

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2004, 06:04:59 ; Search time 59 Seconds
(without alignments)
101.608 Million cell updates/sec

Title: US-10-799-005A-1
Perfect score: 97
Sequence: 1 EPNHLSKIAFKIVSQEPA 19

Scoring table: BLOSUM62
Gapop 10.0 ; Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 6498

Minimum DB seq length: 0
Maximum DB seq length: 19

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_25:**
- 1: sp_archaea:**
 - 2: sp_bacteria:**
 - 3: sp_fungi:**
 - 4: sp_human:**
 - 5: sp_invertebrate:**
 - 6: sp_mammal:**
 - 7: sp_mhc:**
 - 8: sp_organelle:**
 - 9: sp_phase:**
 - 10: sp_plant:**
 - 11: sp_rodent:**
 - 12: sp_virus:**
 - 13: sp_vertebrate:**
 - 14: sp_unclassified:**
 - 15: sp_rviro:**
 - 16: sp_bacteriap:**
 - 17: sp_archaeap:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	26.8	15	11	Q8C9M9 mus musculus
2	26	26.8	19	4	Q9BX20 homo sapien
3	25	25.8	15	4	Q9UMY6 homo sapien
4	25	25.8	15	8	Q9T2J9 pisum sativ
5	25	25.8	16	4	Q8WXZ7
6	25	25.8	18	4	O43834 homo sapien
7	24	24.7	10	2	Q9R5N4 clostridium
8	24	24.7	19	8	Q36925 nicotiana v
9	24	24.7	19	8	Q3G466 brassica ju
10	23	23.7	12	6	Q8M0E4
11	23	23.7	13	8	Q95777
12	23	23.7	14	11	Q06415
13	23	23.7	15	5	Q8ST73 drosophila
14	23	23.7	15	6	Q9TR17 oryctolagus
15	23	23.7	17	8	Q9T2S0 solanum tub
16	23	23.7	18	6	Q9TR05 bos taurus

17	23	23.7	19	10	Q9S8Q0
18	22	22.7	10	2	Q9RSN6
19	22	22.7	11	12	P89269
20	22	22.7	13	2	Q9AIR1
21	22	22.7	13	2	Q9RSN5
22	22	22.7	14	11	Q9QVF3
23	22	22.7	14	11	Q06414
24	22	22.7	15	11	Q9QV62
25	22	22.7	16	5	Q9T2R2
26	22	22.7	17	2	P83329
27	22	22.7	17	2	Q7X3X0
28	22	22.7	17	4	Q8N6B0
29	22	22.7	18	6	Q9SMX1
30	22	22.7	18	6	Q9SMW8
31	22	22.7	18	6	Q9SMX0
32	22	22.7	18	6	Q9SMV5
33	22	22.7	18	6	Q9SMW7
34	22	22.7	18	6	Q9SMW1
35	22	22.7	18	6	Q9SMW5
36	22	22.7	18	6	Q9SMW0
37	22	22.7	18	6	Q9SMW4
38	22	22.7	18	6	Q9SMW6
39	22	22.7	18	6	Q9SMW4
40	22	22.7	18	6	Q9SMW2
41	22	22.7	18	6	Q9SMW8
42	22	22.7	18	6	Q9SMV9
43	22	22.7	18	6	Q9SMV8
44	22	22.7	18	6	Q9SMV2
45	22	22.7	18	6	Q9SMV1

ALIGNMENTS

RESULT 1

Q8C9M9	PRELIMINARY;	PRT;	15 AA.
AC	Q8C9M9;		
DT	01-MAR-2003 (TRENBLrel. 23, Created)		
DT	01-MAR-2003 (TRENBLrel. 23, Last sequence update)		
DT	01-JUN-2003 (TRENBLrel. 24, Last annotation update)		
DE	DEAD/H (Fragment).		
GN	DDX26.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RP	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Thymus;		
RX	MEDLINE=22354683; PubMed=12466851;		
RA	The PANTOM Consortium.		
RA	the RIKEN Genome Exploration Research Group Phase I & II Team;		
RT	"Analysis of the mouse transcriptome based on functional annotation of		
RT	60,770 full-length cDNAs."		
RL	Nature 420:563-573(2002).		
DR	EMBL; AK041758; BAC31055.1; --		
DR	MGI; MGI:1202397; Ddx26.		
FT	NON TER		
SQ	SEQUENCE 15 AA; 1816 MW; CDF3F69D90D63857 CRC64;		

Query Match 26.8%; Score 26; DB 11; Length 15;
Best Local Similarity 80.0%; Pred. No. 1.8e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 NHLNS 7
Db 10 NHINS 14

RESULT 2

Q9BX20	PRELIMINARY;	PRT;	19 AA.
ID	Q9BX20		

```

AC Q9BX20;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE BA470C13.1 (Hypothetical protein KIAA1272) (Fragment).
GN BA470C13.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Peck A.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL161658; CAC36066.1; -.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 19 AA; 2211 MW; 2B88DB0ABFF5A5C0 CRC64;

Query Match 26.8%; Score 26; DB 4; Length 19;
Best Local Similarity 36.4%; Pred. No. 2.3e+03;
Matches 4; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 8 KIAKIVSQSP 18
Db ||:|:|:|
5 KVSFQYLSKGP 15

RESULT 3
Q9UMY6
ID Q9UMY6 PRELIMINARY; PRT; 15 AA.
AC Q9UMY6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE ABO glycosyltransferase (Fragment).
GN ABO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Olsson M.L., Guerreiro J.F., Zago M.A., Chester M.A.;
RT "Molecular analysis of the O alleles at the blood group ABO locus in
RT populations of different ethnic origin reveals novel crossing-over
RT events and point mutations.";
RL Biochem. Biophys. Res. Commun. 234:779-782(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Olsson M.L., Chester M.A.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF170890; AAD51650.1; -.
DR GO; GO:0016740; F:transferase activity; IEA.
KW Transferase.
FT NON TER 1
FT NON-TER 15
SQ SEQUENCE 15 AA; 1762 MW; FB7670E69CC8D18 CRC64;

Query Match 25.8%; Score 25; DB 4; Length 15;
Best Local Similarity 80.0%; Pred. No. 2.6e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPNHL 5
Db ||:|:|:|
5 EPDHL 9

RESULT 4
Q9T2J9
ID Q9T2J9 PRELIMINARY; PRT; 15 AA.
AC Q9T2J9;

```

```

DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Transcript-binding protein (Fragment).
OS Pisum sativum (Garden pea).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE.
RA MEDLINE=94083566; PubMed=8260634;
RA Lakhani S., Khanna N.C., Tewari K.K.;
RT "Nascent transcript-binding protein of the pea chloroplast
RT transcriptionally active chromosome.";
RL Plant Mol. Biol. 23:963-979(1993).
DR PIR; S43321; S43321.
SQ SEQUENCE 15 AA; 1668 MW; 0A778A7774F79D79 CRC64;

Query Match 25.8%; Score 25; DB 8; Length 15;
Best Local Similarity 57.1%; Pred. No. 2.6e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 NHLNSKI 9
Db ||:|:|:|
2 NHNGTI 8

RESULT 5
Q8WXZ7
ID Q8WXZ7 PRELIMINARY; PRT; 16 AA.
AC Q8WXZ7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Alpha-3-galactosylaminyltransferase (EC 2.4.1.40) (Fragment).
GN ABO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21411478; PubMed=11520811;
RA Olsson M.L., Irshaid N.M., Hosseini-Maaf B., Hellberg A., Moulds M.K.,
RA Sareneva H., Chester M.A.;
RT "Genomic analysis of clinical samples with serologic ABO blood
RT grouping discrepancies: identification of 15 novel A and B subgroup
RT alleles.";
RL Blood 98:1585-1593(2001).
DR EMBL; AF324008; AAL37340.1; -.
DR GO; GO:0004380; F:glycoprotein-fucosylgalactoside alpha-N-ace. .; IEA.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. .; IEA.
KW Transferase; Glycosyltransferase.
FT NON TER 1
FT NON-TER 16
SQ SEQUENCE 16 AA; 1849 MW; 67B37670F40CC88D CRC64;

Query Match 25.8%; Score 25; DB 4; Length 16;
Best Local Similarity 80.0%; Pred. No. 2.8e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPNHL 5
Db ||:|:|:|
5 EPDHL 9

RESULT 6
O43834
ID O43834 PRELIMINARY; PRT; 18 AA.
AC O43834;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)

```

DE	RPS19'
GN	RPS19'.

```

AC QSMJE4;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Stem cell tyrosine kinase 1 (Fragment).
GN STK-1.
OS Aotus azarai (Southern owl monkey) (Aotus azarae).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Aotinae; Aotus.
OX NCBI_TaxID=30591;
RN [1]
RP SEQUENCE FROM N.A.
RA Singer S.S., Schmitz J., Schwieck C., Zischler H.;
RT "Molecular Cladistic markers in New World monkey phylogeny.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF489260; AAN01125.1; -
DR GO; GO:0016301; F:Kinase activity; IEA.
KW Kinase.
FT NON_TER 1 1
FT NON_TER 12 12
SQ SEQUENCE 12 AA; 1314 MW; EBBA3924FAB44721 CRC64;

Query Match 23.7%; Score 23; DB 6; Length 12;
Best Local Similarity 71.4%; Pred. No. 4.6e+03;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 EPNHLS 7
Db 2 ETHLS 8

RESULT 11
Q957T7 PRELIMINARY; PRT; 13 AA.
ID Q957T7;
AC Q957T7;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Cytochrome b (Fragment).
OS Trichopsis pumila (pygmy gourami).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Perciformes;
OC Anabantoidae; Belontiidae; Trichopsis.
OX NCBI_TaxID=158454;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang T.Y., Tzeng C.S., Shen S.C.;
RT "Conservation and Phylogeography of Taiwan Paradise Fish, Macropodus opercularis Linnaeus.";
RL Acta Zool. Taiwanica 10:121-134(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Wang T.Y., Tzeng C.S., Shen S.C.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF359378; AAK51436.1; -
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 1 1
SQ SEQUENCE 13 AA; 1443 MW; 55430C67982AAB17 CRC64;

Query Match 23.7%; Score 23; DB 8; Length 13;
Best Local Similarity 62.5%; Pred. No. 5e+03;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 LNSKIAFK 12
Db 6 LEDKILFK 13

RESULT 12
Q06415 PRELIMINARY; PRT; 14 AA.
ID Q06415;
AC Q06415;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE 2X myosin heavy chain (Fragment).
OS Rattus rattus (Black rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10117;
RN [1]
RP SEQUENCE FROM N.A.
RA DeNardi C., Ausoni S., Moretti P., Gorza L., Velleca M.,
RL Buckingham M., Schiaffino S.;
RT "Type 2X Myosin Heavy Chain is coded by a muscle fiber type-specific and developmentally regulated gene.";
RL J. Cell Biol. 123:823-835(1993).
DR EMBL; X72591; CAA51189.1; -
FT NON_TER 1 1
SQ SEQUENCE 14 AA; 1655 MW; F1C0536CD7B6AFA1 CRC64;

Query Match 23.7%; Score 23; DB 11; Length 14;
Best Local Similarity 66.7%; Pred. No. 5.4e+03;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 12 KIVSOE 17
Db 9 KIISEE 14

```

```

RESULT 13
Q8ST73 PRELIMINARY; PRT; 15 AA.
ID Q8ST73;
AC Q8ST73;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Amylase 1 (Fragment).
GN AMY1.
OS Drosophila pseudoobscura (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7237;
RN [1]
RP SEQUENCE FROM N.A.
RA Schaeffer S.W., Goetting-Minesky M.P., Kovacevic M., Peoples J.,
RA Graybill J.L., Miller J.M., Kim K., Nelson J.G., Anderson W.W.;
RT "Evolutionary genomics of inversions in Drosophila pseudoobscura: Modes of selection.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF476197; AAL85805.1; -
DR EMBL; AF476212; AAL85820.1; -
DR EMBL; AF476214; AAL85822.1; -
DR EMBL; AF476215; AAL85823.1; -
DR EMBL; AF476216; AAL85824.1; -
DR EMBL; AF476217; AAL85825.1; -
DR EMBL; AF476219; AAL85827.1; -
DR EMBL; AF476220; AAL85828.1; -
DR FlyBase; FBgn0012691; Dpse\Amy1.
FT NON_TER 1 1
SQ SEQUENCE 15 AA; 1523 MW; CC1DE41B52C5A382 CRC64;

Query Match 23.7%; Score 23; DB 5; Length 15;
Best Local Similarity 50.0%; Pred. No. 5.8e+03;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 HLNSKI 9
Db 10 HVNAKL 15

```


Tue Oct 12 06:45:33 2004

```

RESULT 14
Q9TR17 PRELIMINARY; PRT; 15 AA.
AC Q9TR17; 2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE HISTIOCYTE-secreted-factor (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE.
RX MEDLINE=96179491; PubMed=8603813;
RA Satomi N., Haranaka R., Haranaka K.;
RT "Purification, characterization and anti-tumor activity of a new
RT cytokine, histiocyte-secreted-factor (HSF).";
RL Int. J. Cancer 66:209-213(1996).
SQ SEQUENCE 15 AA; 1689 MW; EC063F77162767E7 CRC64;

Query Match 23.7%; Score 23; DB 6; Length 15;
Best Local Similarity 80.0%; Pred. No. 5.8e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PNHLN 6
Db 11 PNQLN 15

RESULT 15
Q9T2S0 PRELIMINARY; PRT; 17 AA.
AC Q9T2S0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cytochrome-C reductase 55 kDa subunit (EC 1.10.2.2) (Fragment).
OS Solanum tuberosum (Potato).
OG Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE.
RX MEDLINE=94198758; PubMed=7764624;
RA Braun H.P., Kruft V., Schmitz U.K.;
RL Planta 193:99-106(1994).
DR GO; GO:0008121; Fubiquinol-cytochrome-c reductase activity; IEA.
SQ SEQUENCE 17 AA; 1895 MW; 5E2AEB3D17D506B6 CRC64;

Query Match 23.7%; Score 23; DB 8; Length 17;
Best Local Similarity 25.0%; Pred. No. 6.7e+03;
Matches 4; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy 4 HLNSKIAFKIVSQEPA 19
Db 2 HMGSELVQRVAINELA 17

Search completed: October 12, 2004, 06:11:52
Job time : 64 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2004, 06:04:29 ; Search time 15 Seconds
(without alignments)
65.956 Million cell updates/sec

Title: US-10-799-005a-1

Perfect score: 97
Sequence: 1 EPNHLSKIAFKIVSQEPA 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 1082

Minimum DB seq length: 0

Maximum DB seq length: 19

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24	24.7	13	1	PROX_ORYSA
2	24	24.7	16	1	FOR1_MYRGU
3	24	24.7	16	1	FOR2_MYRGU
4	22	22.7	14	1	UC15_MAIZE
5	22	22.7	16	1	CERB_RAT
6	21	21.6	7	1	CHOX_ALCSP
7	21	21.6	13	1	FARB_ASCSU
8	21	21.6	17	1	AU33_LITRA
9	20	20.6	15	1	ARCA_STRP5
10	19	19.6	13	1	BF37_LEUMA
11	19	19.6	15	1	CBPB_PROAT
12	19	19.6	15	1	MP2A_ORYSA
13	19	19.6	15	1	PC20_BRANA
14	19	19.6	18	1	UC03_MAIZE
15	18.5	19.1	19	1	OXL4_OPHHA
16	18	18.6	9	1	FAR1_PANRE
17	18	18.6	8	1	FAR2_PANRE
18	18	18.6	10	1	TEMK_RANTE
19	18	18.6	15	1	FGF1_CANFA
20	18	18.6	15	1	MAOX_CHICK
21	18	18.6	15	1	MCA2_RHOOP
22	18	18.6	17	1	PSBL_SYNVU
23	18	18.6	19	1	DCAM_ACACA
24	18	18.6	19	1	HCY3_PANJA
25	18	18.6	19	1	MIFH_TRIMR
26	17	17.5	8	1	ACI_THUAL
27	17	17.5	10	1	COXQ_RABIT
28	17	17.5	12	1	UH03_RAT
29	17	17.5	13	1	EP65_HUMAN
30	17	17.5	13	1	TEMK_RANTE
31	17	17.5	13	1	TEMK_RANTE
32	17	17.5	15	1	COXI_THUOB
33	17	17.5	15	1	EF1A_MICCR

34	17	17.5	15	1	UP02_METAN
35	17	17.5	16	1	DBH3_RHILE
36	17	17.5	16	1	H5_COTJA
37	17	17.5	17	1	AU32_LITRA
38	17	17.5	17	1	PNOC_PIG
39	17	17.5	17	1	UP37_UPEMJ
40	17	17.5	19	1	UKAI_HUMAN
41	16	16.5	6	1	QVM_LEPDE
42	16	16.5	8	1	RS1_ERWCH
43	16	16.5	9	1	RT33_BOVIN
44	16	16.5	10	1	AH3_PRUSE
45	16	16.5	10	1	ANGT_BOVIN

P83439	metathizium
P80605	rhizobium l
P18638	coturnix co
P82395	litoria ran
P55791	sus scrofa
P82044	uperoleia m
P31940	homo sapien
P42985	leptinotars
P37985	erwinia chr
P82926	bos taurus
P29261	prunus sero
P01017	bos taurus

ALIGNMENTS

RESULT 1

ID	PROX_ORYSA	STANDARD;	PRT;	13 AA.
AC	P83647;			
DT	10-OCT-2003 (Rel. 42, Created)			
DT	10-OCT-2003 (Rel. 42, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Probable profilin LP04 (Fragments).			
OS	Oryza sativa (Rice).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
OC	Ehrhartoideae; Oryzeae; Oryza.			
OX	NCBI_TaxID=4530;			
RN	[1]			
RP	SEQUENCE.			
RC	STRAIN=cv. Indica-IR64; TISSUE=Panicle;			
RA	Hosseini Salekden S.G.; Bennett J.;			
RT	"Proteome analysis of rice panicle.";			
RL	Submitted (JUL-2003) to Swiss-Prot.			
CC	-!- FUNCTION: Binds to actin and affects the structure of the			
CC	cytoskeleton. At high concentrations, profilin prevents the			
CC	polymerization of actin, whereas it enhances it at low			
CC	concentrations. By binding to PIP2, it inhibits the formation of			
CC	IP3 and DG (by similarity).			
CC	-!- SUBUNIT: Occurs in many kinds of cells as a complex with monomeric			
CC	actin in a 1:1 ratio.			
CC	-!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown			
CC	protein is: 4.4, its MW is: 14.0 kDa.			
CC	-!- SIMILARITY: Belongs to the profilin family.			
DR	PROSITE; PS00414; PROFILIN; PARTIAL.			
KW	Actin-binding; Cytoskeleton; Multigene family.			
FT	NON_TER 1			
FT	NON_CONS 5			
FT	NON_TER 13			
FT	NON_TER 13			
SQ	SEQUENCE 13 AA; 1362 MW; OA3022BE052C68B CRC64;			

Query Match 24.7%; Score 24; DB 1; Length 13;
Best Local Similarity 30.0%; Pred. NO. 4.9e+02;
Matches 3; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY	9	IAPKIVSQEP	18
	:	:	:
	:	:	:
Db	4	LAYMVIQGE	13

RESULT 2

ID	FOR1_MYRGU	STANDARD;	PRT;	16 AA.
AC	P81438;			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Formaezin l.			
OS	Myrmecia gulosa (Red bulldog ant).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Vespoidea;			

OC Formicidae; Myrmecini; Myrmecini; Myrmecia.
 RN NCBI_TaxID=36170;
 RP SEQUENCE, AND CARBOHYDRATE-LINKAGE SITE THR-11.
 RC TISSUE=Hemolymph; PubMed=9497332;
 RX MEDLINE=98165787; PubMed=9497332;
 RA Mackintosh J.A., Veal D.A., Beattie A.J., Gooley A.A.;
 RT "Isolation from an ant Myrmecia gulosa of two inducible
 O-glycosylated proline-rich antibacterial peptides.";
 RL J. Biol. Chem. 273:6139-6143 (1998).
 CC -!- FUNCTION: Antibacterial peptide. Has activity against E.coli
 but none against other Gram-negative bacteria and Gram-positive
 bacteria.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- INDUCTION: By bacterial infection.
 CC -!- PTM: O-LINKED GLYCAN CONSISTS OF A GAL-GALNAC DISACCHARIDE, O-
 GLYCOSYLATION IS ESSENTIAL FOR FULL BIOLOGICAL ACTIVITY.
 CC -!- SIMILARITY: TO DROSOPHILA DROSOCIN.
 KW Antibiotic; Glycoprotein; Insect immunity; Hemolymph.
 FT CARBOHYD 11 O-LINKED (GALNAC. . .).
 SQ SEQUENCE 16 AA; 1794 MW; 80CEA3AABEC2E0AE CRC64;
 Query Match 24.7%; Score 24; DB 1; Length 16;
 Best Local Similarity 57.1%; Pred. No. 6.1e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 2 PNHLNSK 8
 Db ||:|:
 3 PNPVNNK 9
 RESULT 3
 ID FOR2 MYRGU STANDARD; PRT; 16 AA.
 AC P81437;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Formacin 2.
 OS Myrmecia gulosa (Red bulldog ant).
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Vespoidea;
 OC Formicidae; Myrmecini; Myrmecia.
 OX NCBI_TaxID=36170;
 RN [1]
 RP SEQUENCE, AND CARBOHYDRATE-LINKAGE SITE THR-11.
 RC TISSUE=Hemolymph;
 RX MEDLINE=98165787; PubMed=9497332;
 RA Mackintosh J.A., Veal D.A., Beattie A.J., Gooley A.A.;
 RT "Isolation from an ant Myrmecia gulosa of two inducible
 O-glycosylated proline-rich antibacterial peptides.";
 RL J. Biol. Chem. 273:6139-6143 (1998).
 CC -!- FUNCTION: Antibacterial peptide. Has activity against E.coli but
 none against other Gram-negative bacteria and Gram-positive
 bacteria.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- INDUCTION: By bacterial infection.
 CC -!- PTM: O-LINKED GLYCAN CONSISTS OF A GAL-GALNAC DISACCHARIDE, O-
 GLYCOSYLATION IS ESSENTIAL FOR FULL BIOLOGICAL ACTIVITY.
 CC -!- SIMILARITY: TO DROSOPHILA DROSOCIN.
 KW Antibiotic; Glycoprotein; Insect immunity; Hemolymph.
 FT CARBOHYD 11 O-LINKED (GALNAC. . .).
 SQ SEQUENCE 16 AA; 1807 MW; 9C3CA3B00BC2E0AE CRC64;
 Query Match 24.7%; Score 24; DB 1; Length 16;
 Best Local Similarity 57.1%; Pred. No. 6.1e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 2 PNHLNSK 8
 Db ||:|:
 3 PNPVNTX 9

RESULT 4
 UC15 MAIZE
 ID UC15 MAIZE STANDARD; PRT; 14 AA.
 AC P80621;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Unknown protein from 2D-page of etiolated coleoptile (Spot 245)
 DE (Fragment).
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
 OX NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Coleoptile;
 RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
 RA Pernollet J.-C., Zivy M., de Vienne D.;
 RT "The maize two dimensional gel protein database: towards an integrated
 genome analysis program.";
 RL Theor. Appl. Genet. 93:997-1005 (1996).
 CC -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
 protein is: 4.8, its MW is: 35.7 kDa.
 DR Maize-2DPAGE; P80621; COLEOPTILE.
 DR MaizeDB; 123947; -. 1
 FT NON_TER 14 14
 FT NON_TER 14 14
 SQ SEQUENCE 14 AA; 1396 MW; C68949275F404CD2 CRC64;
 Query Match 22.7%; Score 22; DB 1; Length 14;
 Best Local Similarity 42.9%; Pred. No. 1.2e+03;
 Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 13 IVSQEPA 19
 Db : :|:
 8 VAAEPPA 14
 RESULT 5
 ID CERR RAT STANDARD; PRT; 16 AA.
 AC P23436; P02682;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cerebellin.
 OS Rattus norvegicus (Rat), and
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116, 9823;
 RN [1]
 RP SEQUENCE.
 RC SPECIES=Rat;
 RA Slemmon J.R., Blacher R., Danho W., Hempstead J.L., Morgan J.I.;
 RT "Isolation and sequencing of two cerebellum-specific peptides.";
 RL Proc. Natl. Acad. Sci. U.S.A. 81:6866-6870 (1984).
 RN [2]
 RP SEQUENCE.
 RC SPECIES=Pig; TISSUE=Brain;
 RX MEDLINE=89341798; PubMed=2760624;
 RA Yiangou Y., Burnet P., Nikou G., Chrysanthou B.J., Bloom S.R.;
 RT "Purification and characterisation of cerebellins from human and
 porcine cerebellum.";
 RL J. Neurochem. 53:886-889 (1989).
 CC -!- FUNCTION: CEREBELLIN EXERTS NEUROMODULATORY FUNCTIONS. DIRECTLY
 STIMULATES NORPINEPRINE RELEASE VIA THE ADENYLATE CYCLASE/PKA-
 DEPENDENT SIGNALING PATHWAY; AND INDIRECTLY ENHANCES
 ADRENOCORTICAL SECRETION IN VIVO, THROUGH A PARACRINE MECHANISM
 INVOLVING MEDULLARY CATECHOLAMINE RELEASE.
 CC -!- TISSUE SPECIFICITY: LOCALIZED IN THE PURKINJE CELLS.
 DR PIR; A03135; CQRT.

DR PIR; PL0124; PL0124.
KW Synaptosome.
SQ SEQUENCE 16 AA; 1633 MW; 3EFA16635343D518 CRC64;
Query Match 22.7%; Score 22; DB 1; Length 16;
Best Local Similarity 33.3%; Pred. No. 1.4e+03;
Matches 3; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
QY 6 NSKIAFKIV 14
Db :|||:
3 SAKVAFSAI 11

RESULT 6
CHOX ALCSP STANDARD; PRT; 7 AA.
AC P16101;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-APR-1990 (Rel. 14, Last annotation update)
DE Choline oxidase (EC 1.1.3.17) (Fragment).
OS Alcaligenes sp.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Alcaligenes.
OX NCBI_TaxID=512;
RN [1]
RP SEQUENCE.
RX MEDLINE=81006769; PubMed=6997283;
RA Ohta-Fukuyama M., Miyake Y., Emi S., Yamano T.;
RT "Identification and properties of the prosthetic group of choline
oxidase from *Alcaligenes* sp.";
RL J. Biochem. 88:197-203(1980).
CC -!- CATALYTIC ACTIVITY: Choline + O(2) = betaine aldehyde + H(2)O(2).
DR PIR; A15398; A15398.
KW Oxidoreductase.
FT NON TER
SQ SEQUENCE 7 AA; 839 MW; 7415B1E457644AC0 CRC64;
Query Match 21.6%; Score 21; DB 1; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 PNH 4
Db :|||:
3 PNH 5

RESULT 7
FARB ASCSU STANDARD; PRT; 13 AA.
AC P43173;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE FMRamide-like neuropeptide AF11.
OS Ascaris suum (pig roundworm) (Ascaris lumbricoides).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC Ascarididae; Ascaris.
OX NCBI_TaxID=6253;
RN [1]
RP SEQUENCE.
RX MEDLINE=95380362; PubMed=7651904;
RA Cowden C., Stretton A.O.W.;
RT "Bright novel FMRamide-like neuropeptides isolated from the nematode
RT *Ascaris suum*.";
RL Peptides 16:491-500(1995).
CC -!- SIMILARITY: Belongs to the FARP (FMRamide related peptide)
CC family.
KW Neuropeptide; Amidation.
FT MOD RES 13 13
SQ SEQUENCE 13 AA; 1495 MW; 9CAEC650D686B05 CRC64;
Query Match 21.6%; Score 21; DB 1; Length 13;

Best Local Similarity 80.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 EPNHL 5
Db :|||:
7 EPNFL 11

RESULT 8
AU33 LITRA STANDARD; PRT; 17 AA.
AC P82396;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Aurein 3.3 [Contains: Aurein 3.3.1].
OS Litoria raniformis (Southern bell frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylidae;
OC Pelodyadinae; Litoria.
OX NCBI_TaxID=116057;
RN [1]
RP SEQUENCE, AND FUNCTION.
RX TISSUE=Skin secretion;
RX MEDLINE=20408845; PubMed=10951191;
RA Rozek T., Wegener K.L., Bowie J.H., Olver I.N., Carver J.A.,
RA Wallace J.C., Tyler M.J.;
RT "The antibiotic and anticancer active aurein peptides from the
RT Australian bell frogs *Litoria aurea* and *Litoria raniformis* the
RT solution structure of aurein 1.2.";
RL Eur. J. Biochem. 267:5330-5341(2000).
CC -!- FUNCTION: Antimicrobial activity against *L. lactis*, *M. luteus*,
CC *P. multocida*, *S. aureus*, *S. epidermidis* and *S. uberis*. Probably acts
CC by disturbing membrane functions with its amphipathic structure.
CC Shows anticancer activity.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.
KW Amphibian defense peptide; Amidation; Antibiotic.
FT PEPTIDE 1 17 AUREIN 3.3.
FT PEPTIDE 3 17 AUREIN 3.3.1.
FT MOD RES 17 17 AMIDATION.
SQ SEQUENCE 17 AA; 1797 MW; F6AC0A471428EAFD CRC64;
Query Match 21.6%; Score 21; DB 1; Length 17;
Best Local Similarity 75.0%; Pred. No. 2.1e+03;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 8 KIAFKIVS 15
Db :|||:
8 KIAGHIVS 15

RESULT 9
ARCA STRPS STANDARD; PRT; 15 AA.
AC P58827;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Arginine deiminase (EC 3.5.3.6) (ADI) (Arginine dihydrolase) (AD)
DE (Streptococcal acid glycoprotein) (Fragment).
GN ARCA OR SAGP.
OS Streptococcus pyogenes (serotype M5).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=160491;
RN [1]
RP SEQUENCE, AND CHARACTERIZATION.
RC STRAIN=Manfredo / Serotype M5.
RX MEDLINE=98298018; PubMed=9632565;
RA Degnan B.A., Palmer J.M., Robson T., Jones C.E.D., Fischer M.,
RA Glanville M., Mellor G.D., Diamond A.G., Kehoe M.A., Goodacre J.A.;
RT "Inhibition of human peripheral blood mononuclear cell proliferation

RT by Streptococcus pyogenes cell extract is associated with arginine
 RL deiminase activity.";
 CC Infect. Immun. 66:3050-3058(1998).
 CC -!- FUNCTION: Antitumor protein. Has a powerful and dose-dependent
 CC inhibitory effect on antigen, superantigen, or mitogen-stimulated
 CC human peripheral blood mononuclear cell (PBMC) proliferation. It
 CC may inhibit cell proliferation by arresting cell cycle and
 CC inducing apoptosis.
 CC -!- CATALYTIC ACTIVITY: L-arginine + H(2)O = L-citrulline + NH(3).
 CC -!- PATHWAY: Arginine degradation via arginine deiminase; first step.
 CC -!- SUBUNIT: Homodimer.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -!- PTM: Glycosylated (By similarity).
 CC -!- SIMILARITY: Belongs to the arginine deiminase family.
 DR HAMAP; MF_00242; -; 1.
 KW Hydrolase; Arginine metabolism; Glycoprotein.
 FT INIT MET 0
 FT NON_TER 15 15
 SQ SEQUENCE 15 AA; 1657 MW; D21150201B00EE46 CRC64;
 Query Match 20.6%; Score 20; DB 1; Length 15;
 Best Local Similarity 50.0%; Pred. No. 2.8e+03;
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 PNHLSKI 9
 DB ||:|:
 5 PIHVYSEI 12
 RESULT 10
 BP37_LEUMA
 ID BP37_LEUMA STANDARD; PRT; 13 AA.
 AC P81754;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Brain protein 37F3.
 OS Leucophaea maderae (Madeira cockroach).
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;
 CC Blaberidae; Leucophaea.
 OX NCBI_TaxID=6988;
 RN [1]
 RP SEQUENCE, AND MASS SPECTROMETRY.
 RC TISSUE=Brain;
 RX MEDLINE=97269266; PubMed=9114447;
 RA Muren J.E., Naessel D.R.;
 RT "Seven tachykinin-related peptides isolated from the brain of the
 RT Madeira cockroach; evidence for tissue-specific expression of
 RT isoforms.";
 RL Peptides 18:7-15(1997).
 CC -!- TISSUE SPECIFICITY: Brain.
 CC -!- MASS SPECTROMETRY: MW=1435.7; METHOD=MALDI.
 SQ SEQUENCE 13 AA; 1436 MW; 8E532C9DE9A7D1B5 CRC64;
 Query Match 19.6%; Score 19; DB 1; Length 13;
 Best Local Similarity 80.0%; Pred. No. 3.5e+03;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 14 VSQEP 18
 DB ||||
 8 VSAEP 12
 RESULT 11
 CBPB_PROAT
 ID CBPB_PROAT STANDARD; PRT; 15 AA.
 AC P19628;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Carboxypeptidase B (EC 3.4.17.2) (Fragment).
 OS Protopterus aethiopicus (Marbled lungfish).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Dipnoi; Lepidosireniiformes; Protopteriidae; Protopterus.
 OX NCBI_TaxID=7886;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Pancreas;
 RX MEDLINE=73025047; PubMed=5079891;
 RA Reek G.R., Neurath H.;
 RT "Isolation and characterization of pancreatic procarboxypeptidase B
 RT and carboxypeptidase B of the African lungfish.";
 RL Biochemistry 11:3947-3955(1972).
 CC -!- CATALYTIC ACTIVITY: Peptidyl-L-lysine (or L-arginine) + H(2)O =
 CC peptide + L-lysine (or L-arginine).
 CC -!- SIMILARITY: Belongs to peptidase family M14.
 DR PIR; A26212; A26212.
 DR MEROPS; M14.003; -.
 DR InterPro; IPR000834; Peptidase M14.
 DR PROSITE; PS00132; CARBOXYPEPT_ZN_1; PARTIAL.
 DR PROSITE; PS00133; CARBOXYPEPT_ZN_2; PARTIAL.
 KW Hydrolase; Carboxypeptidase; Metalloprotease; Zinc; Zymogen.
 FT PROPEP 1 >15
 FT NON_TER 15 15
 SQ SEQUENCE 15 AA; 1749 MW; 124C910D937BED65 CRC64;
 Query Match 19.6%; Score 19; DB 1; Length 15;
 Best Local Similarity 25.0%; Pred. No. 4.1e+03;
 Matches 3; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
 QY 2 PNHLSKIAFKI 13
 DB ||:|:
 4 PRSENGDKVERV 15
 RESULT 12
 MP2A_ORYSA
 ID MP2A_ORYSA STANDARD; PRT; 15 AA.
 AC P83466;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Pollen allergen Ory s 2-A (Fragment).
 OS Oryza sativa (Rice).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 CC Ehrhartoideae; Oryzeae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=cv. Japonica; TISSUE=Pollen;
 RA Kerim T., Imin N., Weinman J.J., Rolfe B.G.;
 RL Submitted (SEP-2002) to Swiss-Prot.
 CC -!- ALLERGEN: Causes an allergic reaction in human. Causes grass
 CC pollen allergy. Binds IGE.
 CC -!- SIMILARITY: Belongs to the expansin family.
 DR InterPro; IPR007112; Expan_endogl.
 DR PROSITE; PS50842; EXPANSIN_EG45; FALSE_NEG.
 KW Allergen.
 FT NON_TER 15 15
 SQ SEQUENCE 15 AA; 1513 MW; 2C65C9FBE3632A1C CRC64;
 Query Match 19.6%; Score 19; DB 1; Length 15;
 Best Local Similarity 28.6%; Pred. No. 4.1e+03;
 Matches 2; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 QY 7 SKIAFKI 13
 DB ||:|:
 1 TEVTFKV 7
 RESULT 13
 PC20_BRANA
 ID PC20_BRANA STANDARD; PRT; 15 AA.
 AC P81096;

DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE 20 kDa pollen coat protein (Fragment).
 OS Brassica napus (Rape).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Brassica.
 OX NCBI_TaxID=3708;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=cv. Topas; TISSUE=Pollen;
 RX MEDLINE=98345939; PubMed=9680961;
 RA Murphy D.J., Ross J.H.;
 RT "Biosynthesis, targeting and processing of oleosin-like proteins,
 RT which are major pollen coat components in Brassica napus.";
 RL Plant J. 13:1-16(1998).
 CC -!- FUNCTION: Major component of the pollen coat.
 CC -!- TISSUE SPECIFICITY: Pollen.
 FT NON_TER 15 15
 SQ SEQUENCE 15 AA; 1756 MW; 4C9B7C58ED18A442 CRC64;

 Query Match 19.6%; Score 19; DB 1; Length 15;
 Best Local Similarity 66.7%; Pred. No. 4.1e+03;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

 Qy 3 NHLNSK 8
 |||||
 Db 9 NHLCHK 14

 RESULT 14
 UC03_MAIZE
 ID UC03_MAIZE STANDARD; PRT; 18 AA.
 AC P80609;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Unknown protein from 2D-page of etiolated coleoptile (Spot 146)
 DE (fragments).
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
 OX NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Coleoptile;
 RA Tourzet P., Ricciardi F., Morin C., Damerval C., Huet J.-C.,
 RA Pernollet J.-C., Zivy M., de Vienne D.;
 RT "The maize two dimensional gel protein database: towards an integrated
 RT genome analysis program.";
 RL Theor. Appl. Genet. 93:997-1005(1996).
 CC -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
 CC protein is: 5.1, its MW is: 29.3 kDa.
 CC -!- CAUTION: The order of the peptides shown is uncertain.
 DR Maize-2DPAGE; P80609; COLEOPTILE.
 DR MaizeDB; 123924; -.
 FT NON_TER 1 1
 FT NON_CONS 9 10
 FT NON_TER 18 18
 SQ SEQUENCE 18 AA; 1938 MW; 6F513BEC35881C0 CRC64;

 Query Match 19.6%; Score 19; DB 1; Length 18;
 Best Local Similarity 42.9%; Pred. No. 4.9e+03;
 Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

 Qy 5 LNSKIAF 11
 |.:|:
 Db 6 LQNKLAY 12

RESULT 15

OXLA_OPHHA
 ID UC03_OPHHA STANDARD; PRT; 19 AA.
 AC P81383;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE L-amino acid oxidase (EC 1.4.3.2) (LAO) (LAO) (Fragment).
 OS Ophiophagus hannah (King cobra) (Naja hannah).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
 OC Elapidae; Elapinae; Ophiophagus.
 OX NCBI_TaxID=8665;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Venom;
 RX MEDLINE=94361525; PubMed=8080286;
 RA Ponnudurai G., Chung M.C.M., Tan N.-H.;
 RT "Purification and properties of the L-amino acid oxidase from Malayan
 RT pit viper (Calloselasma rhodostoma) venom.";
 RL Arch. Biochem. Biophys. 313:373-378(1994).
 RN [2]
 RP SEQUENCE OF 1-15.
 RC TISSUE=Venom;
 RX MEDLINE=97449790; PubMed=9304806;
 RA Ahn M.-Y., Lee B.M., Kim Y.S.;
 RT "Characterization and cytotoxicity of L-amino acid oxidase from the
 RT venom of king cobra (Ophiophagus hannah).";
 RL Int. J. Biochem. Cell Biol. 23:911-919(1997).
 CC -!- FUNCTION: Has cytotoxic activity (By similarity).
 CC -!- CATALYTIC ACTIVITY: An L-amino acid + H(2)O + O(2) = a 2-oxo acid
 CC + NH(3) + H(2)O(2).
 CC -!- COFACTOR: FAD.
 CC -!- SUBUNIT: Homodimer (Probable).
 CC -!- PTM: Glycosylated.
 CC -!- SIMILARITY: BELONGS TO THE FLAVIN MONOAMINE OXIDASE FAMILY.
 CC Strong, to mammalian Fig1.
 KW Oxidoreductase; Flavoprotein; FAD; Glycoprotein; Toxin.
 FT CONFLICT 1 1 H -> S (IN REF. 2).
 FT NON_TER 19 19
 SQ SEQUENCE 19 AA; 2298 MW; DD911ASB414F1427 CRC64;

 Query Match 19.1%; Score 18.5; DB 1; Length 19;
 Best Local Similarity 62.5%; Pred. No. 6.3e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

 Qy 1 EP---NHL 5
 |||||
 Db 11 EPEVENHL 18

Search completed: October 12, 2004, 06:10:47
 Job time : 19 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2004, 06:06:19 ; Search time 22 Seconds
(without alignments)
83.074 Million cell updates/sec

Title: US-10-799-005a-1

Perfect score: 97

Sequence: 1 EPNHLSKIAFKIVSQEPA 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 3435

Minimum DB seq length: 0

Maximum DB seq length: 19

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	26.8	17	2 PH1822	T cell receptor al
2	25	25.8	10	2 S66458	ferredoxin - Rhizo
3	25	25.8	15	2 S43321	RNA-binding protei
4	23	23.7	14	2 A49018	myosin heavy chain
5	23	23.7	17	2 PC2136	zymogen granule me
6	22	22.7	9	2 JP0073	ribosomal protein
7	22	22.7	10	2 F44644	neurotoxin-associa
8	22	22.7	13	2 E39778	lactose phosphotra
9	22	22.7	13	2 G44644	neurotoxin-associa
10	22	22.7	16	2 CORT	cerebellin - rat
11	22	22.7	16	2 P10124	cerebellin - pig
12	21.5	22.2	19	2 A37968	neural surface pro
13	21	21.6	7	2 A15398	choline oxidase (E
14	20	20.6	13	2 JQ2309	hypothetical 1.6K
15	20	20.6	13	2 JQ2319	hypothetical 1.6K
16	20	20.6	14	2 PH0795	T-cell receptor al
17	20	20.6	14	2 PH0776	T-cell receptor al
18	20	20.6	15	2 PA0097	starch phosphoryla
19	20	20.6	16	2 PH0777	T-cell receptor al
20	20	20.6	18	2 C40433	Nada protein - Cya
21	20	20.6	18	2 S70611	30K protein - Enge
22	19.5	20.1	17	2 C43599	hypothetical prote
23	19	19.6	12	2 S71034	potB protein - Sal
24	19	19.6	12	2 F84132	hypothetical prote
25	19	19.6	13	2 PC1008	40K extracellular
26	19	19.6	13	2 S57571	T cell receptor al
27	19	19.6	13	2 A86126	hypothetical prote
28	19	19.6	13	2 E42762	proteasome endopep
29	19	19.6	15	2 P50452	32K protein 3306 -

30 19 19.6 15 2 A26212 carboxypeptidase B
31 19 19.6 19 2 C40634 orf21 5' of eryk -
32 18 18.6 9 2 S70345 amine oxidase (cop
33 18 18.6 10 2 A24407 anicyanin - Paraco
34 18 18.6 10 2 PT0309 Ig heavy chain CRD
35 18 18.6 11 2 E60691 phycobilisome 8X 1
36 18 18.6 12 2 D28551 hypothetical prote
37 18 18.6 12 2 A35855 cytokeinin-binding
38 18 18.6 12 2 S27024 Na+/K+-exchanging
39 18 18.6 14 2 B29743 translation initia
40 18 18.6 14 2 PH0801 T-cell receptor al
41 18 18.6 15 2 PQ0174 stylar glycoprotei
42 18 18.6 15 2 PN0118 hemoglobin beta ch
43 18 18.6 15 2 PS0351 15K protein 5106 -
44 18 18.6 15 2 PA0054 protein QF200017 -
45 18 18.6 15 2 A32971 heparin-binding le

ALIGNMENTS

RESULT 1

PH1822

T cell receptor alpha chain V region (clone 5PBL V alpha 24-5) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999

C:Accession: PH1822

R:Porcetti, S.; Yockey, C.E.; Brenner, M.B.; Balk, S.P.

J. Exp. Med. 178, 1-16, 1993

A:Title: Analysis of T cell antigen receptor (TCR) expression by human peripheral blood

A:Reference number: PH1754; MUID:93301585; PMID:8391057

A:Accession: PH1822

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-17 <POR>

Query Match 26.8%; Score 26; DB 2; Length 17;

Best Local Similarity 60.0%; Pred. No. 4.5e+02;

Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 PNHLSKIAF 11

Db 7 PNSSASKIIF 16

RESULT 2

S66458

ferredoxin - Rhizobium meliloti (fragment)

C:Species: Rhizobium meliloti

C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999

C:Accession: S66458

R:Riedel, K.U.; Jouanneau, Y.; Masepohl, B.; Puehler, A.; Klipp, W.

Eur. J. Biochem. 231, 742-746, 1995

A:Title: A Rhizobium meliloti ferredoxin (FdXN) purified from Escherichia coli donates e

A:Reference number: S66458; MUID:95377307; PMID:7649175

A:Accession: S66458

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-10 <RIE>

C:Genetics:

A:Gene: fdxN

Query Match 25.8%; Score 25; DB 2; Length 10;

Best Local Similarity 50.0%; Pred. No. 3.7e+02;

Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 10 AFKIVSQE 17

Db 1 AFXIIASQ 8

RESULT 3

S43321
 RNA-binding protein - garden pea (fragment)
 C:Species: Pisum sativum (garden pea)
 C>Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 01-Feb-1999
 C:Accession: S43321
 R:Lakhani, S.; Khanna, N.C.; Tewari, K.K.
 Plant Mol. Biol. 23, 963-979, 1993
 A:Title: Nascent transcript-binding protein of the pea chloroplast transcriptionally active
 A:Reference number: S43321; MUID:94083566; PMID:8260634
 A:Accession: S43321
 A:Molecule type: protein
 A:Residues: 1-15 <LAK>
 C:Function:
 A:Description: probably involved in the transcription of chloroplast genes
 C:Keywords: RNA binding

Query Match 25.8%; Score 25; DB 2; Length 15;
 Best Local Similarity 57.1%; Pred. No. 5.9e+02;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 NMLNSKI 9
 |||||
 DB 2 NHINGTI 8

RESULT 4
 A49018
 myosin heavy chain, fast skeletal muscle type X - rat (fragment)
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 13-Aug-1999
 C:Accession: A49018; S32161
 R:DeNardi, C.; Ausoni, S.; Moretti, P.; Gorza, L.; Velleca, M.; Buckingham, M.; Schiaffini, J.; Cell Biol. 123, 823-835, 1993
 A:Title: Type 2X-myosin heavy chain is coded by a muscle fiber type-specific and develop
 A:Reference number: A49018; MUID:94043465; PMID:8227143
 A:Accession: A49018
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-14 <DEN>
 A:Cross-references: GB:X72591; NID:G288645; PIDN:CAA51189.1; PID:G288646
 R:DeNardi, C.; Ausoni, S.; Moretti, P.; Gorza, L.; Velleca, M.; Merlie, J.; Buckingham, M.
 submitted to the EMBL Data Library, March 1993
 A:Description: Type 2X myosin heavy chain is coded by a muscle fiber type-specific and d
 A:Reference number: S32161
 A:Accession: S32161
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-14 <DEN2>
 A:Cross-references: EMBL:X72591; NID:G288645; PIDN:CAA51189.1; PID:G288646
 C:Superfamily: myosin heavy chain; myosin motor domain homology
 C:Keywords: skeletal muscle

Query Match 23.7%; Score 23; DB 2; Length 14;
 Best Local Similarity 66.7%; Pred. No. 1.2e+03;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 12 KIVSOE 17
 |||||
 DB 9 KIISEE 14

RESULT 5
 PC2196
 zymogen granule membrane associated protein, ZAP47 - dog (fragment)
 C:Species: Canis lupus familiaris (dog)
 C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
 C:Accession: PC2196
 R:Fukuoka, S.
 Biosci. Biotechnol. Biochem. 58, 1282-1285, 1994
 A:Title: Analysis of ZAPs, zymogen granule membrane associated proteins, in the regulate
 A:Reference number: PC2195; MUID:94362286; PMID:7765250
 A:Accession: PC2196
 A:Status: preliminary

A:Molecule type: protein
 A:Residues: 1-17 <FUK>

Query Match 23.7%; Score 23; DB 2; Length 17;
 Best Local Similarity 50.0%; Pred. No. 1.5e+03;
 Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EPNHLN 6
 :|||
 DB 5 DENHLN 10

RESULT 6
 JP0073
 ribosomal protein L32 - Leuconostoc mesenteroides (fragment)
 C:Species: Leuconostoc mesenteroides
 C>Date: 10-Mar-1994 #sequence_revision 28-Oct-1994 #text_change 28-Oct-1994
 C:Accession: JP0073
 R:Ochi, K.
 submitted to JIPID, February 1994
 A:Description: Phylogenetic diversity in the genus Bacillus and comparative ribosomal p
 A:Reference number: JP0042
 A:Accession: JP0073
 A:Molecule type: protein
 A:Residues: 1-9 <OCH>
 C:Keywords: protein biosynthesis; ribosome

Query Match 22.7%; Score 22; DB 2; Length 9;
 Best Local Similarity 57.1%; Pred. No. 2.8e+05;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 PNHLNSK 8
 :|||
 DB 3 PSNKNK 9

RESULT 7
 F44644
 neurotoxin-associated protein type B Hn+ 35K chain, band 3a - Clostridium botulinum (fr
 C:Species: Clostridium botulinum
 C>Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 07-Oct-1994
 C:Accession: F44644
 R:Somers, E.; DasGupta, B.R.
 J. Protein Chem. 10, 415-425, 1991
 A:Title: Clostridium botulinum types A, B, C1, and E produce proteins with or without h
 A:Reference number: A44644; MUID:92143938; PMID:1781887
 A:Contents: type B
 A:Accession: F44644
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-10 <SOM>
 A:Note: sequence extracted from NCBI backbone (NCBIP:83787)
 C:Keywords: hemagglutinin

Query Match 22.7%; Score 22; DB 2; Length 10;
 Best Local Similarity 71.4%; Pred. No. 1.2e+03;
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 NMLNSKI 9
 |||||
 DB 3 NILNDKI 9

RESULT 8
 E39778
 lactose phosphotransferase system lacR protein - Lactococcus lactis (fragment)
 C:Species: Lactococcus lactis
 C>Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 11-Jan-2000
 C:Accession: E39778
 R:van Rooijen, R.J.; van Schalkwijk, S.; de Vos, W.M.
 J. Biol. Chem. 266, 7176-7191, 1991
 A:Title: Molecular cloning, characterization, and nucleotide sequence of the tagatose
 A:Reference number: A39778; MUID:9120137; PMID:1901863

A:Accession: E39778
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-13 <VAN>
A:Cross-references: GB:J05748
C:Superfamily: regulatory protein gntR
C:Keywords: DNA binding; transcription regulation

Query Match 22.7%; Score 22; DB 2; Length 13;
Best Local Similarity 50.0%; Pred. No. 1.7e+03;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EPNHNSK 8
| | | | |
Db 3 ESLHWKK 10
| | | | |

RESULT 9
G44644
neurotoxin-associated protein type B Hn+ 35K chain, band 3b - Clostridium botulinum (fra
C:Species: Clostridium botulinum
C:Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 07-Oct-1994
C:Accession: G44644
J. Somers, B.; DasGupta, B.R.
R. Protein Chem. 10, 415-425, 1991
A:Title: Clostridium botulinum types A, B, C1, and E produce proteins with or without he
A:Reference number: A44644; MUID:92143938; PMID:1781887
A:Contents: type B
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-13 <SOW>
A:Note: sequence extracted from NCBI backbone (NCBIP:83785)
C:Keywords: hemagglutinin

Query Match 22.7%; Score 22; DB 2; Length 13;
Best Local Similarity 71.4%; Pred. No. 1.7e+03;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 NNLNSKI 9
| | | | |
Db 6 NILNDKI 12
| | | | |

RESULT 10
CQRT
cerebellin - rat
N:Contains: des-Ser-cerebellin
C:Species: Rattus norvegicus (Norway rat)
C:Date: 20-Jun-2000 #sequence_revision 20-Jun-2000 #text_change 20-Jun-2000
C:Accession: A03135
R. Stiemmon, J.R.; Blacher, R.; Danho, W.; Hempstead, J.L.; Morgan, J.I.
Proc. Natl. Acad. Sci. U.S.A. 81, 6866-6870, 1994
A:Title: Isolation and sequencing of two cerebellum-specific peptides.
A:Reference number: A03135
A:Accession: A03135
A:Molecule type: protein
A:Residues: 1-16 <SLE>
C:Comment: The sequence of des-Ser-cerebellin lacks 1-Ser.
C:Comment: The cerebellin is localized in the Purkinje cells of rat cerebellum.
C:Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal hom
C:Keywords: brain

Query Match 22.7%; Score 22; DB 2; Length 16;
Best Local Similarity 33.3%; Pred. No. 2.1e+03;
Matches 3; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 6 NSKIAFKIV 14
:|:|:|:
Db 3 SAKVAFSAI 11
:|:|:|:

RESULT 11

PL0124
cerebellin - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 18-Aug-2000
C:Accession: PL0124
R. Yiangou, Y.; Burnet, P.; Nikou, G.; Chrysanthou, B.J.; Bloom, S.R.
J. Neurochem. 53, 885-889, 1989
A:Title: Purification and characterisation of cerebellins from human and porcine cerebe
A:Reference number: PL0124; MUID:89341798; PMID:2760624
A:Accession: PL0124
A:Molecule type: protein
A:Residues: 1-16 <YIA>
A:Experimental source: brain
A:Note: the sequences of human and porcine cerebellin are identical to that of the rat
C:Comment: The peptide is identified in two molecular forms having 16 and 15 amino acid
C:Superfamily: unassigned animal peptides

Query Match 22.7%; Score 22; DB 2; Length 16;
Best Local Similarity 33.3%; Pred. No. 2.1e+03;
Matches 3; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 6 NSKIAFKIV 14
:|:|:|:
Db 3 SAKVAFSAI 11
:|:|:|:

RESULT 12
A37968
neural surface protein Bravo - chicken (fragment)
C:Species: Gallus gallus (chicken)
C:Date: 24-Jan-1992 #sequence_revision 24-Jan-1992 #text_change 15-Aug-1997
C:Accession: A37968; A36345
R. de la Rosa, E.J.; Kayyem, J.F.; Roman, J.M.; Stierhof, Y.D.; Dreyer, W.J.; Schwarz, U.
J. Cell Biol. 112, 1049, 1991
A:Reference number: A37968; MUID:91154309; PMID:1999455
A:Contents: erratum
A:Accession: A37968
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-19
R. de la Rosa, E.J.; Kayyem, J.F.; Roman, J.M.; Stierhof, Y.D.; Dreyer, W.J.; Schwarz, U.
J. Cell Biol. 111, 3087-3096, 1990
A:Title: Topologically restricted appearance in the developing chick retinotectal system
A:Reference number: A36345; MUID:91100421; PMID:2269667
A:Accession: A36345
A:Molecule type: protein
A:Residues: 1-7,9-19 <DR2>

Query Match 22.2%; Score 21.5; DB 2; Length 19;
Best Local Similarity 42.9%; Pred. No. 3.1e+03;
Matches 6; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

Qy 5 LNSKIAFKIVSQEP 18
|:|:|:|:|:
Db 5 LDSKLX-EXLSQPP 17
|:|:|:|:|:

RESULT 13
A15398
choline oxidase (EC 1.1.3.17) - Alcaligenes sp. (tentative sequence) (fragment)
C:Species: Alcaligenes sp.
C:Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 31-Mar-2000
C:Accession: A15398
R. Ohta-Fukuyama, M.; Miyake, Y.; Emi, S.; Yamano, T.
J. Biochem. 88, 197-203, 1980
A:Title: Identification and properties of the prosthetic group of choline oxidase from A
A:Reference number: A15398; MUID:81006769; PMID:6997283
A:Accession: A15398
A:Molecule type: protein
A:Residues: 1-7 <OHT>
C:Keywords: oxidoreductase

Query Match 21.6%; Score 21; DB 2; Length 7;

Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PNH 4
|
|
Db 3 PNH 5

RESULT 14

JQ2309
hypothetical 1.6K protein - tomato chloroplast (strain Toko)
C/Species: chloroplast Lycopersicon esculentum (tomato)
C/Date: 10-Mar-1994 #sequence_revision 28-Oct-1994 #text_change 23-Mar-1995
C/Accession: JQ2309
R/Kawagoe, Y.; Kikuta, Y.
Theor. Appl. Genet. 81, 13-20, 1991
A/Title: Chloroplast DNA evolution in potato (Solanum tuberosum L.).
A/Reference number: JQ2306
A/Accession: JQ2309
A/Molecule type: DNA
A/Residues: 1-13 <KAW>
A/Experimental source: strain Toko
C/Genetics:
A/Genome: chloroplast
C/Keywords: chloroplast

Query Match 20.6%; Score 20; DB 2; Length 13;
Best Local Similarity 50.0%; Pred. No. 3.7e+03;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EPNHLN 6
| | |
| | |
Db 3 EKTHIN 8

RESULT 15

JQ2319
hypothetical 1.6K protein - potato chloroplast
C/Species: chloroplast Solanum tuberosum (potato)
C/Date: 10-Mar-1994 #sequence_revision 28-Oct-1994 #text_change 05-Jan-1996
C/Accession: JQ2319; JQ2314
R/Kawagoe, Y.; Kikuta, Y.
Theor. Appl. Genet. 81, 13-20, 1991
A/Title: Chloroplast DNA evolution in potato (Solanum tuberosum L.).
A/Reference number: JQ2306
A/Accession: JQ2319
A/Molecule type: DNA
A/Residues: 1-13 <KW1>
A/Experimental source: cv. W553-4
A/Accession: JQ2314
A/Molecule type: DNA
A/Residues: 1-13 <KW2>
A/Experimental source: cv. 150
C/Genetics:
A/Genome: chloroplast
C/Keywords: chloroplast

Query Match 20.6%; Score 20; DB 2; Length 13;
Best Local Similarity 50.0%; Pred. No. 3.7e+03;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EPNHLN 6
| | |
| | |
Db 3 EKTHIN 8

Search completed: October 12, 2004, 06:12:19
Job time : 26 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2004, 06:00:29 ; Search time 66 Seconds
(without alignments)
81.339 Million cell updates/sec

Title: US-10-799-005A-1

Perfect score: 97

Sequence: 1 EPNHLSKIAFKIVSQEPA 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 490463

Minimum DB seq length: 0
Maximum DB seq length: 19

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	DB ID	Description
1	71	73.2	15	2	AAW04843 Self epit
2	71	73.2	15	2	AAW64815 Desmoglei
3	71	73.2	15	2	AAW78814 Desmoglei
4	71	73.2	15	3	AAAB33625 MHC class
5	71	73.2	15	4	AAAG93721 Human des
6	71	73.2	15	5	AAAO17033 Desmoglei
7	71	73.2	15	6	ABU96577 MHC class
8	32	33.0	15	6	ABP57835 A. margin
9	31	32.0	15	5	ABB04739 Human hep
10	31	32.0	17	2	AAAR21981 hIL-1 bin
11	31	32.0	17	2	AAAR41906 Human IL-
12	31	32.0	17	2	AAAR83473 Anti-infl
13	31	32.0	17	2	AAAR83456 Immunosup
14	30.5	31.4	19	6	ABP71110 E10 prote
15	30	30.9	14	4	AAAG97543 Human pep
16	30	30.9	15	2	AAAY01783 Active fr
17	30	30.9	16	2	AAAR14398 Sclavo pe
18	30	30.9	17	4	AAE08158 Peptide #
19	29	29.9	11	5	ABG31019 Human fib
20	29	29.9	15	2	AAW85386 Helper T-
21	29	29.9	15	2	AAW85174 Helper T-
22	29	29.9	15	3	AAAB13871 L2/ENK1 c
23	29	29.9	16	6	ABP82489 G protein
24	29	29.9	17	6	ABR39983 Rat late
25	28	28.9	10	5	AAE24272 Murine E-

ALIGNMENTS

RESULT 1

AAW04843
ID AAW04843 standard; peptide; 15 AA.

XX AC AAW04843;

XX AC

DT 18-FEB-1997 (first entry)

XX AC

DE Self epitope of desmoglein 3, implicated in autoimmune disease.

XX AC

KW Tolerisation; self-epitope; antigen; autoimmune disease; autoantigen;

KW HLA; human leukocyte antigen; T-cell; thymocyte; pemphigus vulgaris;

KW desmoglein; multiple sclerosis; herpes simplex virus; adenovirus;

KW phosphomannomutase; human papillomavirus; Epstein-Barr virus;

KW DNA polymerase; influenza; haemagglutinin; reovirus; sigma protein.

XX OS Homo sapiens.

XX OS

PN WO9627387-A1.

XX XX

PD 12-SEP-1996.

XX XX

PF 07-MAR-1996; 96WO-US003182.

XX XX

PR 07-MAR-1995; 95US-00400796.

XX XX

FA (HARD) HARVARD COLLEGE.

XX XX

PI Strominger JL, Wucherpfennig KW;

XX XX

DR WPI; 1996-425218/42.

XX XX

PT Pemphigus vulgaris auto-antigens and multiple sclerosis non-self antigens

XX XX

PT - useful in disease treatment, and method for identification of other

XX XX

PS self and non-self antigens implicated in auto-immune disease.

XX XX

CC Claim 1; Page 39; 58pp; English.

XX XX

CC Pharmaceutical preparations for tolerisation to antigens comprise either

XX XX

CC an isolated human non-collagen or non-mysin basic protein (MBP)

XX XX

CC polypeptide which is capable of tolerising an individual to an

XX XX

CC autoantigen; or an isolated human pathogen polypeptide capable of

XX XX

CC tolerising an individual to that polypeptide. In both cases, the

XX XX

CC polypeptide (whether self or non-self) includes an amino acid sequence

ABG71346 Bovine S-
ABG71853 Bovine AK
AAW78190 Human sec
ABB40664 Peptide #
AAM34425 Peptide #
AAM74313 Human bon
AAM61524 Human bra
ABG56112 Human liv
ABG44240 Human pep
AAE08163 Peptide #
AAE08161 Peptide #
AAE08162 Peptide #
AAB73139 Tumour an
ABP55461 Human zin
AAW22728 Membrane
AAW22728 Staphylok
AAE20717 Human Mls
AAE21018 Human Icr
AAY50599 Resin bou
AAY20327 Human mic

26 28 28.9 18 6 ABG71346
27 28 28.9 18 6 ABG71853
28 28 28.9 19 2 AAW78190
29 28 28.9 19 4 ABB40664
30 28 28.9 19 4 AAM34425
31 28 28.9 19 4 AAM74313
32 28 28.9 19 4 AAM61524
33 28 28.9 19 4 ABG56112
34 28 28.9 19 5 ABG44240
35 27 27.8 8 4 AAE08163
36 27 27.8 8 4 AAE08161
37 27 27.8 8 4 AAE08162
38 27 27.8 10 4 AAB73139
39 27 27.8 15 5 ABP55461
40 27 27.8 17 2 AAW22728
41 27 27.8 17 4 AAW22728
42 27 27.8 17 5 AAE20717
43 27 27.8 17 5 AAE21018
44 27 27.8 19 2 AAY50599
45 26 26.8 5 2 AAY20327

CC 3 protein (amino acids 190-204) and is implicated as a self epitope in
CC pemphigus vulgaris. Peptides derived from the human desmoglein protein
CC are described in AAW04841-47
XX
SQ Sequence 15 AA;

Query Match 73.2%; Score 71; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVSQEPA 19
|||||
DB 1 LNSKIAFKIVSQEPA 15

RESULT 2
AAW64815
ID AAW64815 standard; peptide; 15 AA.
XX
AC AAW64815;
XX
DT 29-SEP-1998 (first entry)
XX
DE Desmoglein-3 190-204.
XX
KW Desmoglein; DG; gene therapy; pemphigus vulgaris; microparticle;
KW autoantigen; autoimmune disease; MHC.
XX
OS Homo sapiens.
XX
PN US5783567-A.
XX
PD 21-JUL-1998.
XX
PF 22-JAN-1997; 97US-00787547.
XX
PR 22-JAN-1997; 97US-00787547.
XX
PA (PANG-) PANGAEA PHARM INC.
XX
PI Langer RS, Hedley ML, Curley JM;
XX
DR WPI; 1998-427077/36.
XX
PT Microparticle encapsulated nucleic acids - for recombinant expression of
PT proteins e.g. in gene therapy.
XX
PS Disclosure; Col 4; 42pp; English.

CC The patent describes a new preparation of microparticles each comprising
CC a polymeric matrix and a nucleic acid. The polymeric matrix consists of
CC one or more synthetic polymers having a solubility in water of less than
CC 1 mg/l (e.g. poly-lactic-co-glycolic acid); and at least 90% of the
CC microparticles have a diameter of less than 100 microns. The
CC microparticles are useful for the delivery of nucleic acids to phagocytic
CC cells. In one embodiment the microparticles are less than 20 microns in
CC diameter and the nucleic acid (preferably in closed circular form)
CC includes an expression control sequence operatively linked to a coding
CC sequence, where the expression product of the coding sequence is a
CC polypeptide having a length and a sequence which permits it to bind to an
CC MHC class I or II molecule. The expression product is thus an effective
CC stimulator of an immune response in mammals. The present sequence, an
CC antigenic portion of desmoglein 3, is an example of an MHC class II
CC peptide which can be expressed by the nucleic acid. It is associated with
CC pemphigus vulgaris
XX

Query Match 73.2%; Score 71; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVSQEPA 19

DB 1 LNSKIAFKIVSQEPA 15
|||||

RESULT 3
AAW78814
ID AAW78814 standard; peptide; 15 AA.
XX
AC AAW78814;
XX
DT 17-NOV-1998 (first entry)
XX
DE Desmoglein 3 protein fragment 190-204.
XX

KW Microparticle; delivery; polymeric matrix; autoantigen; tumour antigen;
KW class II associated peptide; pathogen; gene therapy; genetic disease;
KW infection; downregulation; immune response.

OS Homo sapiens.
XX
PN WO9831398-A1.
XX
PD 23-JUL-1998.
XX
PF 22-JAN-1998; 98WO-US001499.
XX
PR 22-JAN-1997; 97US-00787547.
XX
PR 06-JAN-1998; 98US-00003253.
XX
PA (PANG-) PANGAEA PHARM INC.
XX
PI Hedley ML, Curley JM, Langer RS, Lunsford LB;
XX
DR WPI; 1998-427556/36.
XX
PT New preparations of microparticles - comprising a synthetic polymer
PT matrix and nucleic acid comprising an expression vector for use in gene
PT therapy.
XX
PS Disclosure; Page 8; 101pp; English.

CC A microparticle preparation (MP) has been developed, consisting of
CC microparticles having a diameter of less than 100 mu m. The MP comprises:
CC (a) a polymeric matrix (PM) consisting of one or more synthetic polymers
CC having a solubility in water of less than 1 mg/l; and (b) an expression
CC vector selected from RNA molecules (at least 50% of which are closed
CC circles) or circular plasmid DNA (at least 50% of which are supercoiled).
CC Also described is a MP of at most 20 microns in diameter, comprising: (a)
CC a PM; and (b) a NAM comprising an expression control sequence operatively
CC linked to a coding sequence, where the coding sequence encodes an
CC expression product selected from: (i) a polypeptide at least 7 amino
CC acids in length, having a sequence identical to the sequence of: (i) a
CC fragment of a naturally-occurring mammalian protein; or (ii) a fragment
CC of a naturally-occurring protein from an infectious agent which infects a
CC mammal; (2) a peptide having a length and sequence which permits it to
CC bind to an MHC class I or II molecule; and (3) the polypeptide or the
CC peptide linked to a trafficking sequence. AAW69763 to AAW69765, and
CC AAW78793 to AAW78897 are peptide fragments for use in the present
CC invention. The MPs are highly effective vehicles for the delivery of
CC polynucleotides into phagocytic cells. They can be used for gene therapy,
CC e.g. for treating genetic diseases, infections or tumours or for
CC downregulating an immune response
XX

SQ Sequence 15 AA;

Query Match 73.2%; Score 71; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVSQEPA 19
|||||
DB 1 LNSKIAFKIVSQEPA 15

PN WO200206316-A2.
 PD 24-JAN-2002.
 XX
 XX
 PF 16-JUL-2001; 2001WO-US022263.
 XX
 PR 14-JUL-2000; 2000US-0218381P.
 PR 18-AUG-2000; 2000US-0226382P.
 PR 06-OCT-2000; 2000US-0238380P.
 PR 29-DEC-2000; 2000US-0258764P.
 PR 14-JUN-2001; 2001US-0298317P.
 XX
 XX (ZYCO-) ZYCOS INC.
 XX
 XX Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;
 XX WPI; 2002-195801/25.
 DR
 XX Novel nucleic acid encoding fusion protein comprising alpha-melanocyte
 PT stimulating hormone concatamer or its analog, for treating inflammatory
 PT or autoimmune disorders.
 XX
 XX Disclosure; Page 26; 89pp; English.
 PS
 XX The present invention relates to a nucleic acid comprising a sequence
 CC encoding a fusion polypeptide having an alpha-melanocyte stimulating
 CC hormone (MSH) concatamer. The sequences are useful for treating an
 CC individual suffering from, or at risk of, a disorder of the immune system
 CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid
 CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact
 CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,
 CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The
 CC present sequence is a peptide described in the exemplification of the
 CC invention
 XX
 SQ Sequence 15 AA;
 Query Match 73.2%; Score 71; DB 5; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.4e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 LNSKIAFKIVSQEPA 19
 DB 1 LNSKIAFKIVSQEPA 15
 RESULT 7
 ABU96577
 ID ABU96577 standard; peptide; 15 AA.
 XX
 AC ABU96577;
 XX
 DT 12-AUG-2003 (first entry)
 XX
 DE MHC class II associated desmoglein 3 peptide 190-204.
 XX
 KW Microparticle; microsphere; polynucleotide delivery; phagocytic cell;
 KW tumour; viral infection; bacterial infection; fungal infection;
 KW protozoan infection; gene therapy; major histocompatibility complex;
 KW MHC class II.
 XX
 XX Unidentified.
 OS
 XX US2002182258-A1.
 PN
 XX
 PD 05-DEC-2002.
 XX
 XX 18-JUL-2001; 2001US-00909460.
 PF
 XX 22-JAN-1997; 97US-0035983P.
 PR 06-JAN-1998; 98US-00003253.
 PR 22-JAN-1998; 98WO-US0001499.
 PR 11-MAR-1999; 99US-00266463.

PR 27-MAY-1999; 99US-00321346.
 XX
 PA (ZYCO-) ZYCOS INC.
 XX
 PI Lunsford LB, Putnam D, Hedley ML;
 XX WPI; 2003-438782/41.
 DR
 XX Microparticles, useful as vehicles for delivery of polynucleotides to
 PT phagocytic cells, comprises polymeric matrix, lipid, and nucleic acid
 PT molecule.
 XX
 XX Disclosure; Page 3; 37pp; English.
 PS
 XX The invention relates to a microparticle (microsphere) less than 20
 CC microns in diameter that comprises: (1) a polymeric matrix; (2) a lipid;
 CC and (3) a nucleic acid molecule. The microparticle is not encapsulated in
 CC a liposome and the microparticle does not comprise a cell. The
 CC microparticles are used as vehicles for the delivery of polynucleotides
 CC into phagocytic cells. The microparticles can be used to express antigens
 CC to treat tumour cells or viral, bacterial, fungal or protozoan
 CC infections. The microparticles can be made without adversely affecting
 CC nucleic acid integrity. The present sequence represents the amino acid
 CC sequence of a major histocompatibility complex, MHC, class II associated
 CC peptide
 XX
 SQ Sequence 15 AA;
 Query Match 73.2%; Score 71; DB 6; Length 15;
 Best Local Similarity 100.0%; Pred. No. 1.4e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5 LNSKIAFKIVSQEPA 19
 DB 1 LNSKIAFKIVSQEPA 15
 RESULT 8
 ABP57835
 ID ABP57835 standard; protein; 15 AA.
 XX
 AC ABP57835;
 XX
 DT 03-FEB-2003 (first entry)
 XX
 DE A. marginale Msp-2 fragment.
 XX
 KW Ehrlichiae; Rickettsiae; pathogen; antimicrobial; vaccine;
 KW immune response; Msp-2.
 XX
 OS Anaplasma marginale.
 XX
 XX WO200283724-A1.
 PN
 XX 24-OCT-2002.
 PD
 XX 11-APR-2002; 2002WO-AU000468.
 PF
 XX 12-APR-2001; 2001AU-00004400.
 PR 10-SEP-2001; 2001AU-00007597.
 PR 01-MAR-2002; 2002AU-00000861.
 XX
 XX (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
 PA
 XX Riding GA, Hope MA, Willadsen P;
 PI WPI; 2003-093001/08.
 DR
 XX New vaccine, useful for the manufacture of a medicament for raising an
 PT immune response against Ehrlichiae and/or Rickettsiae pathogens.
 PT
 XX Example 1; Page 32; 102pp; English.
 PS
 XX

CC The invention relates to a novel vaccine comprising at least one
 CC polypeptide which raises an immune response against Ehrlichiae and/or
 CC Rickettsiae pathogens when administered to a subject. The vaccine of the
 CC invention has antimicrobial activity. The vaccine is useful for the
 CC manufacture of a medicament for raising an immune response against
 CC Ehrlichiae and/or Rickettsiae pathogens or for treating or preventing
 CC Ehrlichiae or Rickettsiae infection in a subject. The present sequence
 CC represents a fragment of the A. marginale Msp-2 polypeptide
 XX
 SQ Sequence 15 AA;

Query Match 33.0%; Score 32; DB 6; Length 15;
 Best Local Similarity 40.0%; Pred. No. 1.2e+02;
 Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 4 HNSKIAFKI 13
 ||| |
 Db 5 HINPKFAYR 14
 ||| |

RESULT 9
 ABB04739
 ID ABB04739 standard; peptide; 15 AA.
 XX
 AC ABB04739;
 XX
 XX
 DT 11-MAR-2002 (first entry)
 XX
 XX Human heparan sulfate 3-O-sulfotransferase 19 peptide SEQ ID NO:7.

XX Human; heparan sulfate 3-O-sulfotransferase 19; haemorrhagic disease;
 KW thrombus embolism; haemopathy; myocardial infarction; tumour;
 KW inflammation; immunological disease; HIV infection.
 XX

OS Homo sapiens.

XX

PN CN1311305-A.

XX

PD 05-SEP-2001.

XX

PF 02-MAR-2000; 2000CN-00111794.

XX

PR 02-MAR-2000; 2000CN-00111794.

XX

PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.

XX

PI Mao Y, Xie Y;

XX

DR WPI; 2002-049921/07.

XX

PT New polypeptide-heparan sulfate 3-O-sulfotransferase 19 and

PT polynucleotide for coding such polypeptide.

XX

PS Example 6; Page 19 (Disclosure); 34pp; Chinese.

XX

CC The present invention describes human heparan sulfate 3-O-

CC sulfotransferase 19 protein (I). The present invention also describes a

CC method of applying (I) in the treatment of various diseases, such as

CC haemorrhagic diseases, thrombus embolism, other haemopathy, myocardial

CC infarction, various tumours, inflammation, immunological diseases and HIV

CC infection. The present invention also describes the agonist resisting (I)

CC and its treatment effect. The present sequence represents the N-terminal

CC peptide of (I) which is used in an example from the present invention

XX

SQ Sequence 15 AA;

Query Match 32.0%; Score 31; DB 5; Length 15;

Best Local Similarity 42.9%; Pred. No. 1.8e+02;

Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 5 LNSKIAFKIVSQEP 18
 :||| |
 Db 1 MNSSIKLLIIVREP 14
 :||| |

RESULT 10
 AAR21981

ID AAR21981 standard; peptide; 17 AA.

XX

AC AAR21981;

XX

DT 25-MAR-2003 (revised)

DT 25-JUN-1992 (first entry)

XX

DE hIL-1 binding peptide.

XX

XX Human Interleukin-1; inflammatory diseases; rheumatoid arthritis;

KW osteoporosis.

XX

OS Synthetic.

XX

PN EP474141-A.

XX

PD 11-MAR-1992.

XX

PF 30-AUG-1991; 91EP-00114678.

XX

PR 03-SEP-1990; 90JP-00233571.

PR

24-NOV-1990; 90JP-00320000.

PR

26-NOV-1990; 90JP-00324956.

XX

XX (KURS) KURARAY CO LTD.

PA

Suda T, Abe E, Tanihara M, Fujiwara C;

PI

WPI; 1992-081703/11.

XX

PT New peptide(s) with specific binding activity for human interleukin-1 -

PT for treating inflammatory diseases caused by excess interleukin-1 prodn.

PT

e.g. rheumatoid arthritis and osteoporosis.

XX

XX Claim 3; Page 6; 9pp; English.

PS

The peptide and its salts has high binding activity for human interleukin

XX

-1. The peptides can be made by solid phase synthesis. IL-1 is an

CC

important factor in immune response and is involved in inflammatory

CC

diseases. The peptide inhibits IL-1 activity without binding to the IL-1

CC

receptor and is used in treating diseases arising from IL-1, e.g.

CC

rheumatoid arthritis and osteoporosis. See also AAR21978-R21993. (Updated

CC

on 25-MAR-2003 to correct PA field.)

XX

SQ Sequence 17 AA;

Query Match 32.0%; Score 31; DB 2; Length 17;

Best Local Similarity 50.0%; Pred. No. 2.1e+02;

Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 5 LNSKIAFKIVSQEP 18
 |::| |
 Db 4 LRIKISAKFVNEP 17
 |::| |

RESULT 11

AAR41906

ID AAR41906 standard; peptide; 17 AA.

XX

AC AAR41906;

XX

XX 25-MAR-2003 (revised)

DT 21-APR-1994 (first entry)

XX

XX Human IL-1 binding peptide.

DE

Interleukin 1; IL-1; binding; osteoporosis.

XX

XX Synthetic.

OS

```

XX PN JP05229958-A.
XX
XX PD 07-SEP-1993.
XX PF 10-MAR-1992; 92JP-00087802.
XX PR 26-DEC-1991; 91JP-00359538.
XX PA (KURS ) KURARAY CO LTD.
XX FA (MIZU/) MIZUSHIMA H.
XX
XX DR WPI; 1993-317396/40.
XX
XX PT Agent to treat osteoporosis - contg. peptide serine and cysteine, which
XX PT binds to human interleukin 1.
XX PS Example; Page 6; 8pp; Japanese.
XX
XX CC The peptide is used as an agent for the treatment of osteoporosis.
XX CC (Updated on 25-MAR-2003 to correct PR field.)
XX
XX SQ Sequence 17 AA;
    Query Match      32.0%; Score 31; DB 2; Length 17;
    Best Local Similarity 50.0%; Pred. No. 2.1e+02;
    Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVSQEP 18
   | | | | |
Db 4 LRIKISAKFVNEP 17

RESULT 12
AAR83473
ID AAR83473 standard; peptide; 17 AA.
XX
XX AC AAR83473;
XX
XX DT 20-DEC-1995 (first entry)
XX
XX DE Anti-inflammatory interleukin-1 binding peptide #3.
XX
XX KW Anti-inflammatory agent; human interleukin-1; hIL-1; binding peptide;
XX KW rheumatoid arthritis; nephritis; periodontosis; dermatitis;
XX KW ulcerative colitis; arteriosclerosis; pulmonary silicosis.
XX
XX OS Synthetic.
XX
XX PN JP07025783-A.
XX
XX PD 27-JAN-1995.
XX
XX PF 26-DEC-1991; 91JP-00359537.
XX
XX PR 26-DEC-1991; 91JP-00359537.
XX
XX PA (KURS ) KURARAY CO LTD.
XX FA (MIZU/) MIZUSHIMA Y.
XX
XX DR WPI; 1995-101811/14.
XX
XX PT Antiinflammatory agents contg. human interleukin-1 binding peptide -
XX PT useful for treating e.g. rheumatoid arthritis, nephritis or dermatitis.
XX
XX PS Disclosure; Page 3; 8pp; Japanese.
XX
XX CC AAR83471-R83486 are specific examples of a generic formula (see AAR83470)
XX CC for interleukin-1 binding peptides. The peptides are useful as anti-
XX CC inflammatory agents, e.g. for treatment of inflammatory diseases such as
XX CC rheumatoid arthritis, periodontosis, nephritis, dermatitis, ulcerative
XX CC colitis, arteriosclerosis and pulmonary silicosis
XX

```

```

SQ Sequence 17 AA;
    Query Match      32.0%; Score 31; DB 2; Length 17;
    Best Local Similarity 50.0%; Pred. No. 2.1e+02;
    Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVSQEP 18
   | | | | |
Db 4 LRIKISAKFVNEP 17

RESULT 13
AAR83456
ID AAR83456 standard; peptide; 17 AA.
XX
XX AC AAR83456;
XX
XX DT 20-DEC-1995 (first entry)
XX
XX DE Immunosuppressive IL-1-binding peptide #3.
XX
XX KW Immunosuppression; interleukin-1; IL-1; allergic asthma;
XX KW multiple sclerosis; systemic lupus erythematosus; SLE;
XX KW transplant rejection.
XX
XX OS Synthetic.
XX
XX PN JP07025782-A.
XX
XX PD 27-JAN-1995.
XX
XX PF 26-DEC-1991; 91JP-00359539.
XX
XX PR 26-DEC-1991; 91JP-00359539.
XX
XX PA (KURS ) KURARAY CO LTD.
XX FA (MIZU/) MIZUSHIMA Y.
XX
XX DR WPI; 1995-101810/14.
XX
XX PT Immunosuppressive agent contg. interleukin-1 binding peptide - used in
XX PT treatment of transplant rejection and IL-1 related diseases, e.g.
XX PT allergic asthma, multiple sclerosis and SLE.
XX
XX PS Disclosure; Page 3; 7pp; Japanese.
XX
XX CC AAR83454-R83469 are specific examples of a generic formula (see AAR83453)
XX CC for interleukin-1 binding peptides. The peptides, which have low
XX CC toxicity, are useful as immunosuppressive agents, e.g. for treatment of
XX CC transplant rejection and IL-1 related diseases
XX
SQ Sequence 17 AA;
    Query Match      32.0%; Score 31; DB 2; Length 17;
    Best Local Similarity 50.0%; Pred. No. 2.1e+02;
    Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVSQEP 18
   | | | | |
Db 4 LRIKISAKFVNEP 17

RESULT 14
ABP71110
ID ABP71110 standard; peptide; 19 AA.
XX
XX AC ABP71110;
XX
XX DT 14-APR-2003 (first entry)
XX
XX DE E10 protein CARD region fragment.
XX
XX KW BTF3; cell death; apoptosis; basic transcription factor; cytostatic;

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2004, 06:12:00 ; Search time 21 Seconds
(without alignments)
46.709 Million cell updates/sec

Title: US-10-799-005A-1

Perfect score: 97
Sequence: 1 EPNHLSKIAFKIVSQEPA 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 166134

Minimum DB seq length: 0
Maximum DB seq length: 19

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	71	73.2	15	1	US-08-787-547-24
2	71	73.2	15	2	US-08-400-756-3
3	31	32.0	12	2	US-08-637-759B-448
4	31	32.0	12	3	US-08-871-355A-448
5	31	32.0	12	4	US-09-201-945-448
6	31	32.0	17	1	US-08-239-854-6
7	30	30.9	17	3	US-08-160-604-62
8	29	29.9	15	4	US-09-009-953-76
9	28	28.9	18	4	US-09-268-480-21
10	28	28.9	18	4	US-09-268-480-28
11	27	27.8	8	3	US-08-160-604-65
12	27	27.8	8	3	US-08-160-604-66
13	27	27.8	8	3	US-08-160-604-67
14	27	27.8	19	4	US-09-442-989-30
15	26	26.8	9	3	US-09-518-046-101
16	26	26.8	15	1	US-08-204-656B-16
17	26	26.8	15	1	US-08-470-702-16
18	26	26.8	15	1	US-08-467-831-16
19	26	26.8	19	3	US-09-308-935-15
20	25	25.8	10	4	US-09-490-702B-57
21	25	25.8	12	4	US-08-634-332A-21
22	25	25.8	12	4	US-08-634-332A-22
23	25	25.8	13	1	US-08-548-540-152
24	25	25.8	13	1	US-08-191-338A-29
25	25	25.8	13	1	US-08-191-338A-30
26	25	25.8	13	1	US-08-191-338A-31
27	25	25.8	13	5	PCT-US96-09809-152

28 25 25.8 15 1 US-08-049-783-11 Sequence 11, Appl
29 25 25.8 15 1 US-08-158-232-22 Sequence 22, Appl
30 25 25.8 15 1 US-08-304-626-22 Sequence 22, Appl
31 25 25.8 15 1 US-08-316-301A-23 Sequence 23, Appl
32 25 25.8 15 2 US-08-611-928-22 Sequence 22, Appl
33 25 25.8 15 3 US-09-224-024-5 Sequence 5, Appl
34 25 25.8 15 3 US-09-173-891-22 Sequence 22, Appl
35 25 25.8 15 3 US-09-076-137-23 Sequence 23, Appl
36 25 25.8 15 4 US-09-009-953-23 Sequence 23, Appl
37 25 25.8 15 4 US-09-009-953-46 Sequence 46, Appl
38 25 25.8 15 4 US-09-009-953-134 Sequence 134, Appl
39 25 25.8 15 4 US-09-009-953-223 Sequence 223, Appl
40 25 25.8 15 4 US-09-009-953-224 Sequence 224, Appl
41 25 25.8 15 4 US-09-311-784A-418 Sequence 418, Appl
42 25 25.8 15 4 US-09-738-363-23 Sequence 23, Appl
43 25 25.8 15 5 PCT-US92-03624-23 Sequence 23, Appl
44 25 25.8 15 5 PCT-US94-07902-5 Sequence 5, Appl
45 25 25.8 16 1 US-07-752-101A-15 Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-08-787-547-24
; Sequence 24, Application US/08787547
; Patent No. 5783567
; GENERAL INFORMATION:
; APPLICANT: Hedley, Mary Lynne
; APPLICANT: Curley, Joanne M.
; APPLICANT: Langer, Robert S.
; TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY
; TITLE OF INVENTION: OF NUCLEIC ACID
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/787,547
; FILING DATE: 22-JAN-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 08191/003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-787-547-24

Query Match 73.2%; Score 71; DB 1; Length 15;
Best Local Similarity 100.0%; Pred.No. 3.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Tue Oct 12 06:45:32 2004

us-10-799-005a-1.closed.ra1

QY 5 LNSKIAFKIVSQEPA 19
Db 1 LNSKIAFKIVSQEPA 15

RESULT 2
US-08-400-796-3
; Sequence 3, Application US/08400796
; Patent No. 5874531
; GENERAL INFORMATION:
; APPLICANT: STROMINGER, JACK L.
; APPLICANT: WUCHERPENNIG, KAI
; TITLE OF INVENTION: IDENTIFICATION OF SELF AND NON-SELF
; TITLE OF INVENTION: ANTIGENS IMPLICATED IN AUTOIMMUNE DISEASE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02210

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/400,796
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: TWOMEY, MICHAEL J.
; REGISTRATION NUMBER: 38,349
; REFERENCE/DOCKET NUMBER: H0498/7015
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: HOMO SAPIENS

US-08-400-796-3
Query Match 73.2%; Score 71; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 LNSKIAFKIVSQEPA 19
Db 1 LNSKIAFKIVSQEPA 15

RESULT 3
US-08-637-759B-448
; Sequence 448, Application US/08637759B
; Patent No. 5876931
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia

; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/637,759B
; FILING DATE: 03-MAY-1996
; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 448:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-637-759B-448

Query Match 32.0%; Score 31; DB 2; Length 12;
Best Local Similarity 50.0%; Pred. No. 33;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
QY 4 HLNSKIAFKIVS 15
Db 1 YANNIIAFQVVS 12

RESULT 4
US-08-871-355A-448
; Sequence 448, Application US/08871355A
; Patent No. 6015669
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/871,355A
; FILING DATE: 09-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284

REFERENCE/DOCKET NUMBER: RIMS 101 CON
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 448:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-871-355A-448

Query Match 32.0%; Score 31; DB 3; Length 12;
Best Local Similarity 50.0%; Pred. No. 33;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 4 HNSKIAFKIVS 15
: | : | | : | |
Db 1 YANNIIAQVVS 12

RESULT 5
US-09-201-945-448
; Sequence 448, Application US/09201945
; Patent No. 6342215
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/201,945
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/637,759
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RIMS 101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 448:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-09-201-945-448

Query Match 32.0%; Score 31; DB 4; Length 12;
Best Local Similarity 50.0%; Pred. No. 33;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 4 HNSKIAFKIVS 15
: | : | | : | |
Db 1 YANNIIAQVVS 12

RESULT 6
US-08-239-854-6
; Sequence 6, Application US/08239854
; Patent No. 5624900
; GENERAL INFORMATION:
; APPLICANT: Suda, Tatsuo
; APPLICANT: Abe, Etsuko
; APPLICANT: Tanihara, Masao
; APPLICANT: Fujiwara, Chie
; TITLE OF INVENTION: Peptide or Its Salts
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/239,854
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/752,426
; FILING DATE:
; APPLICATION NUMBER: JP 233,571/90
; FILING DATE: 03-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 320,000/90
; FILING DATE: 24-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 324,956/90
; FILING DATE: 26-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5624900man P.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 363-279-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 521-4500
; TELEFAX: (703) 486-2347
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-239-854-6

Query Match 32.0%; Score 31; DB 1; Length 17;
Best Local Similarity 50.0%; Pred. No. 50;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVSQEP 18
: | : | | : | |
Db 4 LRIKISAKFVNEP 17

RESULT 7
US-08-160-604-62
; Sequence 62, Application US/08160604
; Patent No. 6232522
; GENERAL INFORMATION:

APPLICANT: Harley, John
APPLICANT: James, Judith A.
APPLICANT: Scofield, R. H.
TITLE OF INVENTION: PEPTIDE INDUCTION OF AUTOIMMUNITY AND CLINICAL SYMPTOMATOLOGY
NUMBER OF SEQUENCES: 127
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Babst
STREET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-4530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/160,604
FILING DATE: 30-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/867,819
FILING DATE: 13-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/648,205
FILING DATE: 31-JAN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/472,947
FILING DATE: 31-JAN-1990
ATTORNEY/AGENT INFORMATION:
NAME: Babst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMRF114CIP(3)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-815-6508
TELEFAX: (404)-815-6555
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-160-604-62

Query Match 30.9%; Score 30; DB 3; Length 17;
Best Local Similarity 60.0%; Pred. No. 76;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 NHLNLSKIAFK 12
||| ||| :
Db 5 NHLKSKEVWK 14

RESULT 8
US-09-009-953-76
Sequence 76, Application US/09009953
Patent No. 6413517
GENERAL INFORMATION:
APPLICANT: Sette, Alessandro
TITLE OF INVENTION: Identification of Broadly
Reactive DR Restricted Epitopes
NUMBER OF SEQUENCES: 274
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA

ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/009,953
FILING DATE: 21-Jan-1998
CLASSIFICATION: <UNKNOWN>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/036,713
FILING DATE: 23-JAN-1997
APPLICATION NUMBER: US 60/037,432
FILING DATE: 07-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 018623-011520US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
TELEX: <UNKNOWN>
INFORMATION FOR SEQ ID NO: 76:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 76:
US-09-009-953-76

Query Match 29.9%; Score 29; DB 4; Length 15;
Best Local Similarity 66.7%; Pred. No. 98;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 11 FKIVSQEPA 19
||| ||| :
Db 4 FKIGSSDPA 12

RESULT 9
US-09-268-480-21
Sequence 21, Application US/09268480A
Patent No. 6451528
GENERAL INFORMATION:
APPLICANT: Carr, Daniel W
APPLICANT: Vijayaraghavan, Srinivasan
TITLE OF INVENTION: Sperm-specific AKAP Protein Genes and Uses
FILE REFERENCE: 98700
CURRENT APPLICATION NUMBER: US/09/268,480A
CURRENT FILING DATE: 1999-03-16
NUMBER OF SEQ ID NOS: 33
SOFTWARE: Patent in Ver. 2.1
SEQ ID NO 21
LENGTH: 18
TYPE: PRT
ORGANISM: Bos sp.
US-09-268-480-21

Query Match 28.9%; Score 28; DB 4; Length 18;
Best Local Similarity 55.6%; Pred. No. 1.8e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 8 KIAFKIVSQ 16
: ||| :
Db 4 RAAEQIISQ 12

RESULT 10
US-09-268-480-28
Sequence 28, Application US/09268480A


```

; Patent No. 6451528
; GENERAL INFORMATION:
; APPLICANT: Carr, Daniel W
; APPLICANT: Vijayaraghavan, Srinivasan
; TITLE OF INVENTION: Sperm-specific AKAP Protein Genes and Uses
; FILE REFERENCE: 98700
; CURRENT APPLICATION NUMBER: US/09/268,480A
; CURRENT FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Bos sp.
US-09-268-480-28

```

Query Match 28.9%; Score 28; DB 4; Length 18;
Best Local Similarity 50.0%; Pred. NO. 1.8e+02;
Matches 6; Conservative 2; Mismatches 4; Indels

Qy 5 LNSKIAFKIVSQ 16
| | | : | | |
Db 1 LEEKVAAALVSO 12

RESULT 11
US-08-160-604-65
; Sequence 65, Application US/08160604
; Patent No. 6232522
; GENERAL INFORMATION:
; APPLICANT: Harley, John
; APPLICANT: James, Judith A.
; APPLICANT: Scofield, R. H.
; TITLE OF INVENTION: PEPTIDE INDUCTION OF AUTOIMMUNITY AND CLINICAL SYMPTOMATOLOGY
; NUMBER OF SEQUENCES: 127
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/160,604
; FILING DATE: 30-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/867,819
; FILING DATE: 13-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/648,205
; FILING DATE: 31-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/472,947
; FILING DATE: 31-JAN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMRF114CIP(3)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)-815-6508
; TELEFAX: (404)-815-6555
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```

; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
US-08-160-604-65

Query Match      27.8%      Score 27;   DB 3;   Length 8;
Best Local Similarity 83.3%
Matches 5;   Conservative 0; Mismatches 1; Indels 1;

Qy      3   NHLNSK 8
        ||||
Db      3   NHLKSK 8

```

RESULT 12
US-08-160-604-66
; Sequence 66, Application US/08160604
; Patent No. 623252
; GENERAL INFORMATION:
; APPLICANT: Harley, John
; APPLICANT: James, Judith A.
; APPLICANT: Scofield, R. H.
; TITLE OF INVENTION: PEPTIDE INDUCTION OF AUTOIMMUNITY AND CLINICAL SYMPTOMATOLOGY
; NUMBER OF SEQUENCES: 127
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/160,604
; FILING DATE: 30-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/867,819
; FILING DATE: 13-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/648,205
; FILING DATE: 31-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/472,947
; FILING DATE: 31-JAN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMRF114CIP(3)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)-815-6508
; TELEFAX: (404)-815-6555
; INFORMATION FOR SEQ ID NO: 66:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; US-08-160-604-66

Query Match 27.8%; Score 27; DB 3; Length 8;
Best Local Similarity 83.3%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 NHLNSK 8
||| ||
Db 2 NHLKSK 7

RESULT 13
US-08-160-604-67
; Sequence 67, Application US/08160604
; Patent No. 6232522
; GENERAL INFORMATION:
; APPLICANT: Harley, John
; APPLICANT: James, Judith A.
; APPLICANT: Scofield, R. H.
; TITLE OF INVENTION: PEPTIDE INDUCTION OF AUTOIMMUNITY AND CLINICAL SYMPTOMATOLOGY
; NUMBER OF SEQUENCES: 127
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/160,604
; FILING DATE: 30-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/867,819
; FILING DATE: 13-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/648,205
; FILING DATE: 31-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/472,947
; FILING DATE: 31-JAN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMRF114CIP(3)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)-815-6508
; TELEFAX: (404)-815-6555
; INFORMATION FOR SEQ ID NO: 67:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
US-08-160-604-67

Query Match 27.8%; Score 27; DB 3; Length 8;
Best Local Similarity 83.3%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 NHLNSK 8
||| ||
Db 1 NHLKSK 6

RESULT 14
US-09-442-989-30
; Sequence 30, Application US/09442989
; Patent No. 6569993
; GENERAL INFORMATION:

; APPLICANT: Sledeski, Adam W.
; APPLICANT: Mencil, James J.
; TITLE OF INVENTION: PROCESS FOR THE PREPARATION OF RESIN-BOUND CYCLIC
; TITLE OF INVENTION: PEPTIDES
; FILE REFERENCE: A3113B-US
; CURRENT APPLICATION NUMBER: US/09/442,989
; CURRENT FILING DATE: 1999-11-18
; EARLIER APPLICATION NUMBER: 60/081,897
; EARLIER FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 30
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)
; OTHER INFORMATION: FMOC-Lys(BOC)
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (2)
; OTHER INFORMATION: His(Trt)
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (4)
; OTHER INFORMATION: Asn(Trt)
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (5)
; OTHER INFORMATION: Ser(OtBu)
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (6)..(10)
; OTHER INFORMATION: "Synthetic Peptide: The side chains of Lys at
; OTHER INFORMATION: position 6 and Asp at position 10 are linked by an
; OTHER INFORMATION: amide bond."
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (7)
; OTHER INFORMATION: Glu(OtBu)
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (8)
; OTHER INFORMATION: Arg(PMC)
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (6)..(10)
; OTHER INFORMATION: "Synthetic Peptide: The side chains of Lys at
; OTHER INFORMATION: position 6 and Asp at position 10 are linked by an
; OTHER INFORMATION: amide bond."
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (11)
; OTHER INFORMATION: Trp(BOC)
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (13)
; OTHER INFORMATION: Arg(PMC)
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (14)
; OTHER INFORMATION: Lys(BOC)
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (17)
; OTHER INFORMATION: Gln(Trt)
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (16)
; OTHER INFORMATION: Asp(OtBu)
; FEATURE:
; NAME/KEY: PEPTIDE

; LOCATION: (6)..(10)
; OTHER INFORMATION: "Synthetic Peptide: The side chains of Lys at
; OTHER INFORMATION: position 6 and Asp at position 10 are linked by an
; OTHER INFORMATION: amide bond."
US-09-442-989-30

Query Match 27.8%; Score 27; DB 4; Length 19;
Best Local Similarity 100.0%; Pred.No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 HLNSK 8
Db 2 HLNSK 6

RESULT 15

US-09-518-046-101
; Sequence 101, Application US/09518046
; Patent No. 6294663
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Transmembrane Serine Protease Overexpressed
; TITLE OF INVENTION: in Ovarian Carcinoma and Uses Thereof
; FILE REFERENCE: D6192CIP
; CURRENT APPLICATION NUMBER: US/09/518,046
; CURRENT FILING DATE: 2000-03-02
; EARLIER APPLICATION NUMBER: 09/261,416
; EARLIER FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 153
; SEQ ID NO 101
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: 283...291
; OTHER INFORMATION: TADG-12 peptide
US-09-518-046-101

Query Match 26.8%; Score 26; DB 3; Length 9;
Best Local Similarity 62.5%; Pred.No. 3e+05;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PNLNSKI 9
Db 2 PSHLVEKI 9

Search completed: October 12, 2004, 06:17:38
Job time : 22 secs

This Page Blank (uspto)

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	71	73.2	15	9	US-09-909-460-24	Sequence 24, Appl	
2	71	73.2	15	12	US-09-872-836-24	Sequence 24, Appl	
3	30.5	31.4	19	14	US-10-153-344-10	Sequence 10, Appl	
4	29	29.9	15	12	US-10-103-395-76	Sequence 76, Appl	
5	29	29.9	15	14	US-10-186-867-30	Sequence 30, Appl	
6	29	29.9	16	14	US-10-225-567A-1162	Sequence 1162, Appl	
7	29	29.9	18	14	US-10-029-386-33316	Sequence 33316, Appl	
8	28	28.9	18	14	US-10-043-487-406	Sequence 406, Appl	
9	28	28.9	18	14	US-10-244-488-21	Sequence 21, Appl	
10	28	28.9	18	14	US-10-244-488-28	Sequence 28, Appl	
11	28	28.9	19	9	US-09-864-761-44631	Sequence 44631, A	
12	27	27.8	14	16	US-10-712-425-1268	Sequence 1268, Appl	
13	27	27.8	17	9	US-09-946-175-6	Sequence 6, Appl	
14	26	26.8	9	12	US-10-334-726-187	Sequence 187, Appl	
15	26	26.8	9	14	US-10-357-175-101	Sequence 101, Appl	

RESULT 2
US-09-872-836-24
; Sequence 24, Application US/09872836
; Publication No. US20040142475A1
; GENERAL INFORMATION:
; APPLICANT: Barman, Shikha P.

; APPLICANT: McKeever, Una
; APPLICANT: Hedley, Mary Lynne
; TITLE OF INVENTION: DELIVERY SYSTEMS FOR BIOACTIVE AGENTS
; FILE REFERENCE: 08191-018001
; CURRENT APPLICATION NUMBER: US/09/872,836
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/208,830
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-872-836-24

Query Match 73.2%; Score 71; DB 12; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVSQEPA 19
|||||
Db 1 LNSKIAFKIVSQEPA 15

RESULT 3
US-10-153-344-10
; Sequence 10, Application US/10153344
; Publication No. US20030004124A1
; GENERAL INFORMATION:
; APPLICANT: ROTHMAN, JOEL
; APPLICANT: BLOSS, TIM
; TITLE OF INVENTION: BTF3: AN INHIBITOR OF APOPTOSIS
; FILE REFERENCE: 407T-300410US
; CURRENT APPLICATION NUMBER: US/10/153,344
; CURRENT FILING DATE: 2002-08-27
; PRIOR APPLICATION NUMBER: US 60/292,559
; PRIOR FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-153-344-10

Query Match 31.4%; Score 30.5; DB 14; Length 19;
Best Local Similarity 47.1%; Pred. No. 4.2e+02;
Matches 8; Conservative 4; Mismatches 2; Indels 3; Gaps 1;

QY 3 NHIKSKIAFKIVSQEPA 19
:|:|:|
Db 2 DHLRAK---KILSREDA 15

RESULT 4
US-10-103-395-76
; Sequence 76, Application US/10103395
; Publication No. US20020160019A1
; GENERAL INFORMATION:
; APPLICANT: EPIMUNE, Inc.
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; TITLE OF INVENTION: IDENTIFICATION OF BROADLY REACTIVE DR
; TITLE OF INVENTION: RESTRICTED EPITOPES
; FILE REFERENCE: 39963-20016.01
; CURRENT APPLICATION NUMBER: US/10/103,395
; CURRENT FILING DATE: 2003-01-03
; PRIOR APPLICATION NUMBER: US 09/009,953
; PRIOR FILING DATE: 1998-01-21
; PRIOR APPLICATION NUMBER: PCT/US98/01373

; PRIOR FILING DATE: 1998-01-23
; PRIOR APPLICATION NUMBER: US 60/036,713
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 60/037,432
; PRIOR FILING DATE: 1997-02-07
; NUMBER OF SEQ ID NOS: 274
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-103-395-76

Query Match 29.9%; Score 29; DB 12; Length 15;
Best Local Similarity 66.7%; Pred. No. 5.8e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 11 FKIVSQEPA 19
|||
Db 4 FKIGSSDPA 12

RESULT 5
US-10-186-867-30
; Sequence 30, Application US/10186867
; Publication No. US20030100508A1
; GENERAL INFORMATION:
; APPLICANT: Simon-Haldi, Maryline
; APPLICANT: Schachner, Melitta
; APPLICANT: Neuburger, Timothy
; APPLICANT: Uri, Herzberg
; TITLE OF INVENTION: CARBOHYDRATE EPI TOPE MIMIC COMPOUNDS AND USES THEREOF
; FILE REFERENCE: 1094-1-008NCIP
; CURRENT APPLICATION NUMBER: US/10/186,867
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: US/09/511,956
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/121,327
; PRIOR FILING DATE: 1999-02-24
; PRIOR APPLICATION NUMBER: 60/155,492
; PRIOR FILING DATE: 1999-09-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 15
; TYPE: PRT
; ORGANISM: ml3 library
US-10-186-867-30

Query Match 29.9%; Score 29; DB 14; Length 15;
Best Local Similarity 45.5%; Pred. No. 5.8e+02;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVS 15
|:|:|
Db 2 LHTLLRFRVS 12

RESULT 6
US-10-225-567A-1162
; Sequence 1162, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burner, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTOR
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19

```

; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1162
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-225-567A-1162

Query Match      29.9%; Score 29; DB 14; Length 16;
Best Local Similarity 60.0%; Pred. No. 6.2e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 4 HLNSKIAPKI 13
   |||||:
Db 5 HLESKISKRI 14

RESULT 7
US-10-029-386-33316
; Sequence 33316, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AROMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33316
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC000051.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.44
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.43
US-10-029-386-33316

Query Match      29.9%; Score 29; DB 14; Length 18;
Best Local Similarity 83.3%; Pred. No. 7.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 HLNSKI 9
   |||||
Db 9 HLNSKL 14

RESULT 8
US-10-043-487-406
; Sequence 406, Application US/10043487
; Publication No. US20030055220A1
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS
; APPLICANT: Pierre, LEGRAIN
; TITLE OF INVENTION: Protein-protein interactions between Shigella Flexneri polypeptides
; FILE REFERENCE: B4778A
; CURRENT APPLICATION NUMBER: US/10/043,487
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/261,130
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 561
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 406
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Shigella Flexneri
; FEATURE:
; NAME/KEY: misc_feature

```

```

; LOCATION: (4)..(4)
; OTHER INFORMATION: MISC_FEATURE
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7)..(8)
; OTHER INFORMATION: MISC_FEATURE
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (13)..(13)
; OTHER INFORMATION: MISC_FEATURE
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (15)..(15)
; OTHER INFORMATION: MISC_FEATURE
US-10-043-487-406

```

```

Query Match      28.9%; Score 28; DB 14; Length 18;
Best Local Similarity 35.7%; Pred. No. 1e+03;
Matches 5; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

```

```

Qy 5 LNSKIAPKIVSQEP 18
   |::|::|
Db 5 LRXXVAIEXLQEP 18

```

```

RESULT 9
US-10-244-488-21
; Sequence 21, Application US/10244488
; Publication No. US20030104495A1
; GENERAL INFORMATION:
; APPLICANT: Carr, Daniel W
; APPLICANT: Vijayaraghavan, Srinivasan
; TITLE OF INVENTION: Sperm-specific AKAP Protein Genes and Uses
; FILE REFERENCE: 98700
; CURRENT APPLICATION NUMBER: US/10/244,488
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US/09/268,480
; PRIOR FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Bos sp.
US-10-244-488-21

```

```

Query Match      28.9%; Score 28; DB 14; Length 18;
Best Local Similarity 55.6%; Pred. No. 1e+03;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 8 KIAFKIVSQ 16
   :|||:
Db 4 RAAAFQIISQ 12

```

```

RESULT 10
US-10-244-488-28
; Sequence 28, Application US/10244488
; Publication No. US20030104495A1
; GENERAL INFORMATION:
; APPLICANT: Carr, Daniel W
; APPLICANT: Vijayaraghavan, Srinivasan
; TITLE OF INVENTION: Sperm-specific AKAP Protein Genes and Uses
; FILE REFERENCE: 98700
; CURRENT APPLICATION NUMBER: US/10/244,488
; CURRENT FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US/09/268,480
; PRIOR FILING DATE: 1999-03-16
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 18
; TYPE: PRT

```

; ORGANISM: Bos sp.
US-10-244-488-28

Query Match 28.9%; Score 28; DB 14; Length 18;
Best Local Similarity 50.0%; Pred. No. 1e+03;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 LNSKIAPKIVSQ 16
| :| :|||
Db 1 LEEKVAALVSQ 12

RESULT 11

US-09-864-761-44631
; Sequence 44631, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecwica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 44631
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC018528.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1

; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.75
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.84
US-09-864-761-44631

Query Match 28.9%; Score 28; DB 9; Length 19;
Best Local Similarity 50.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 6 NSKIAPKIVSQE 17
| :| :|||
Db 6 NYKLSVKFSQE 17

RESULT 12

US-10-712-425-1268
; Sequence 1268, Application US/10712425
; Publication No. US20040180380A1
; GENERAL INFORMATION:
; APPLICANT: LEE, FRANK D.
; APPLICANT: MENG, XUN
; APPLICANT: LIVINGSTON, DAVID
; TITLE OF INVENTION: PROTEOME EPITOPE TAGS AND METHODS OF USE THEREOF IN PROTEIN
; FILE REFERENCE: MODIFICATION ANALYSIS
; CURRENT APPLICATION NUMBER: US/10/712,425
; CURRENT FILING DATE: 2003-11-13
; PRIOR APPLICATION NUMBER: 60/379,626
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/393,137
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/393,197
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/393,211
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/393,223
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/393,233
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/393,235
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/393,280
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/430,948
; PRIOR FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: 60/433,319
; PRIOR FILING DATE: 2002-12-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1268
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Illustrative peptide having high affinity for its antigen
US-10-712-425-1268

Query Match 27.8%; Score 27; DB 16; Length 14;
Best Local Similarity 55.6%; Pred. No. 1.2e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PNHLNSKIA 10
| :| :|||
Db 4 PEDLNQKLA 12

RESULT 13

US-09-946-175-6
; Sequence 6, Application US/09946175
; Patent No. US20020106671A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert

; TITLE OF INVENTION: IDENTIFICATION OF A CAPACITATIVE CALCIUM
; TITLE OF INVENTION: CHANNEL IN ANTIGEN PRESENTING CELLS AND USES THEREOF

; FILE REFERENCE: Mlsn1(310800)
; CURRENT APPLICATION NUMBER: US/09/946,175

; CURRENT FILING DATE: 2001-09-05

; NUMBER OF SEQ ID NOS: 49

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 6

; LENGTH: 17

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-946-175-6

Query Match 27.8%; Score 27; DB 9; Length 17;

Best Local Similarity 47.1%; Pred. No. 1.4e+03;

Matches 8; Conservative 2; Mismatches 5; Indels 2; Gaps 1;

Qy 1 EPNHLSKIAPKIVSQE 17

Db 3 EPGKLSQKI--KWLQE 17

RESULT 14

US-10-334-726-187

; Sequence 187, Application US/10334726

; Publication No. US20030211521A1

; GENERAL INFORMATION:

; APPLICANT: TAYLOR-PAPADIMITROU, JOYCE

; TITLE OF INVENTION: BREAST CANCER ANTIGEN

; FILE REFERENCE: 1090-36

; CURRENT APPLICATION NUMBER: US/10/334,726

; CURRENT FILING DATE: 2003-01-02

; PRIOR APPLICATION NUMBER: US/09/645,446

; PRIOR FILING DATE: 2000-08-25

; PRIOR APPLICATION NUMBER: PCT/GB99/00866

; PRIOR FILING DATE: 1999-03-19

; PRIOR APPLICATION NUMBER: GB 9805877.9

; PRIOR FILING DATE: 1998-09-20

; NUMBER OF SEQ ID NOS: 324

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 187

; LENGTH: 9

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: predicted

; OTHER INFORMATION: peptide

US-10-334-726-187

Query Match 26.8%; Score 26; DB 12; Length 9;

Best Local Similarity 83.3%; Pred. No. 1.2e+06;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PNHLSNLS 7

Db 1 PVLHLS 6

RESULT 15

US-10-357-175-101

; Sequence 101, Application US/10357175

; Publication No. US20030170707A1

; GENERAL INFORMATION:

; APPLICANT: O'Brien, Timothy J.

; TITLE OF INVENTION: Transmembrane Serine Protease Overexpressed

; TITLE OF INVENTION: in Ovarian Carcinoma and Uses Thereof

; FILE REFERENCE: D6192CIP/D/CIP

; CURRENT APPLICATION NUMBER: US/10/357,175

; CURRENT FILING DATE: 2003-02-03

; PRIOR APPLICATION NUMBER: 09/650,371

; PRIOR FILING DATE: 2000-08-28

; NUMBER OF SEQ ID NOS: 158

; SEQ ID NO 101

; LENGTH: 9

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; LOCATION: 283...291

; OTHER INFORMATION: TADG-12 peptide

US-10-357-175-101

Query Match 26.8%; Score 26; DB 14; Length 9;

Best Local Similarity 62.5%; Pred. No. 1.2e+06;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PNHLSNLS 9

Db 2 PSHLVEKI 9

Search completed: October 12, 2004, 06:16:22

Job time : 70 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2004, 05:56:44 ; Search time 23 Seconds
(without alignments)
79.463 Million cell updates/sec

Title: US-10-799-005A-1

Perfect score: 97

Sequence: 1 EPNHLSKIAFKIVSQEPA 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*

1: piri:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	97	100.0	999	1 IJHUG3	desmoglein 3 precu
2	81	83.5	1043	1 IJBOG1	desmoglein 1 precu
3	81	83.5	1049	1 IJHUG1	desmoglein 1 precu
4	71	73.2	1117	2 S38673	desmoglein 2 - hum
5	52	53.6	533	2 S57904	virR49 protein - S
6	52	53.6	533	2 JC5040	positive regulator
7	52	53.6	906	1 IJHUCN	cadherin 2 precurs
8	49	50.5	877	1 IJBOCN	N-cadherin precurs
9	49	50.5	906	1 IJMSCN	N-cadherin precurs
10	48	49.5	829	2 I46536	Ksp-cadherin - rab
11	47	48.5	317	2 B96963	uncharacterized co
12	46	47.4	783	2 I50116	N-cadherin precurs
13	45	46.4	620	2 A53731	translation initia
14	45	46.4	884	1 IJMSCN	E-cadherin precurs
15	45	46.4	884	2 S34438	uvomorulin - mouse
16	45	46.4	1321	2 S27337	multidrug resistan
17	45	46.4	1321	2 T23476	hypothetical prote
18	44	45.4	347	2 D69373	immunogenic protei
19	44	45.4	512	2 T38422	probable chromatin
20	44	45.4	905	1 IJXLC1	N-cadherin 1 precu
21	44	45.4	906	1 IJXLC2	N-cadherin 2 precu
22	44	45.4	1114	2 T25083	hypothetical prote
23	44	45.4	1158	2 T25082	hypothetical prote
24	43	44.3	307	2 T40240	dimethylase - fiis
25	43	44.3	315	2 T43249	rRNA (adenine-N6,N
26	43	44.3	480	2 T18675	hypothetical prote
27	43	44.3	725	1 E64211	virulence-associat
28	43	44.3	809	1 IJBOOD	desmocollin 2b pre
29	43	44.3	827	2 A53954	LI-cadherin precu

30	43	44.3	863	1 IJBODC	desmocollin 2a pre
31	43	44.3	912	1 IJCHCN	N-cadherin precurs
32	42	43.3	388	2 T09489	hypothetical prote
33	42	43.3	479	2 H72155	CLL protein - vari
34	42	43.3	479	2 T28480	hypothetical prote
35	42	43.3	479	2 T30794	hypothetical prote
36	42	43.3	479	2 E42508	EIL protein - vacc
37	42	43.3	479	2 D36841	EIL protein - vari
38	42	43.3	487	2 F84751	hypothetical prote
39	42	43.3	533	2 E86412	hypothetical prote
40	42	43.3	761	1 IJBODE	desmocollin 1a - b
41	42	43.3	770	2 B48910	desmocollin 1b pre
42	42	43.3	824	2 A48910	desmocollin 1a pre
43	42	43.3	839	1 IJBODF	desmocollin 1b pre
44	42	43.3	840	2 I37281	Desla precursor -
45	42	43.3	873	2 F86426	95.1K hypothetical

ALIGNMENTS

RESULT 1

IJHUG3

desmoglein 3 precursor - human

N;Alternate names: pemphigus vulgaris antigen

C;Species: Homo sapiens (man)

C;Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 22-Jun-1999

C;Accession: A41088

R;Amagai, M.; Klaus-Kovtun, V.; Stanley, J.R.

Cell 67, 869-877, 1991

A;Title: Autoantibodies against a novel epithelial cadherin in pemphigus vulgaris, a dis

A;Reference number: A41088; MUID:92069753; PMID:1720352

A;Accession: A41088

A;Molecule type: mRNA

A;Residues: 1-999 <AMA>

A;Cross-references: GB:M76482; NID:G190751; PIDN:AAA60230.1; PID:G190752

C;Genetics:

A;Gene: GDB:DSG3

A;Cross-references: GDB:I34030; OMIM:169615

A;Map position: 18q12.1-18q12.2

C;Superfamily: cadherin; cadherin repeat homology

C;Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane pro

F;1-23/Domain: signal sequence #status predicted <SIG>

F;24-49/Domain: propeptide #status predicted <PRO>

F;50-999/Product: desmoglein homolog #status predicted <MAT>

F;50-615/Domain: extracellular #status predicted <EXT>

F;52-157/Domain: cadherin repeat homology <CR1>

F;160-267/Domain: cadherin repeat homology <CR2>

F;270-383/Domain: cadherin repeat homology <CR3>

F;390-495/Domain: cadherin repeat homology <CR4>

F;496-598/Domain: cadherin repeat homology <CR5>

F;616-639/Domain: transmembrane #status predicted <TM>

F;640-999/Domain: intracellular #status predicted <INT>

F;910-938/Domain: desmoglein repeat <DGI>

F;937-966/Domain: desmoglein repeat <DG2>

F;110,180,545/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 97; DB 1; Length 999;

Best Local Similarity 100.0%; Pred. No. 1.8e-08;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFKIVSQEPA 19

Db 186 EPNHLSKIAFKIVSQEPA 204

RESULT 2

IJBOG1

desmoglein 1 precursor - bovine

N;Alternate names: desmoglein BDGM

C;Species: Bos primigenius taurus (cattle)

C;Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 22-Jun-1999

C;Accession: S14603; A38872; A37785; S38721; A48173; S24412

R; Koch, P.J.; Goldschmidt, M.D.; Zimbelmann, R.; Franke, W.W.
submitted to the EMBL Data Library, March 1991
A; Description: Complete sequence of the desmoglein precursor and evidence for the existence of a desmoglein-like protein in the desmosomal junctions.
A; Reference number: S14603
A; Accession: S14603
A; Molecule type: mRNA
A; Residues: 1-1043 <KOC>
A; Cross-references: EMBL:X58466; NID:G306; PIDN:CAA41380.1; PID:G307
R; Koch, P.J.; Goldschmidt, M.D.; Walsh, M.J.; Zimbelmann, R.; Franke, W.W.
Eur. J. Cell Biol. 55, 200-208, 1991
A; Title: Complete amino acid sequence of the epidermal desmoglein precursor polypeptide
A; Reference number: A38872; MUID:92037656; PMID:1935985
A; Accession: A38872
A; Molecule type: mRNA
A; Residues: 1-87; 98-1043 <XO2>
A; Cross-references: GB:S64268; GB:S64270
R; Goodwin, L.; Hill, J.E.; Raynor, K.; Raszi, L.; Manabe, M.; Cowin, P.
Biochem. Biophys. Res. Commun. 173, 1224-1230, 1990
A; Title: Desmoglein shows extensive homology to the cadherin family of cell adhesion molecules
A; Reference number: A37785; MUID:91097553; PMID:1702628
A; Accession: A37785
A; Molecule type: mRNA
A; Residues: 44-123, 'V', 125-493 <GOO>
A; Cross-references: GB:M58165; NID:G162966; PIDN:AAA62709.1; PID:G552318
R; Zimbelmann, R.
submitted to the EMBL Data Library, February 1991
A; Reference number: S38721
A; Accession: S38721
A; Molecule type: mRNA
A; Residues: 44-1043 <ZIM>
A; Cross-references: EMBL:X57784; NID:G436061; PIDN:CAA40930.1; PID:G436062
R; Koch, P.J.; Walsh, M.J.; Schmelz, M.; Goldschmidt, M.D.; Zimbelmann, R.; Franke, W.W.
Eur. J. Cell Biol. 53, 1-12, 1990
A; Title: Identification of desmoglein, a constitutive desmosomal glycoprotein, as a member of the desmosomal glycoprotein family
A; Reference number: A48173; MUID:91168965; PMID:1706270
A; Accession: A48173
A; Molecule type: mRNA
A; Residues: 44-1001, 'AQPSPAT' <KOS>
A; Cross-references: GB:X57784
A; Note: this sequence has been revised in references A38872 and S38721
C; Genetics:
A; Gene: DSG1
C; Superfamily: cadherin; cadherin repeat homology
C; Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane protein
F; 1-23/Domain: signal sequence #status predicted <SIG>
F; 24-49/Domain: propeptide #status predicted <PRO>
F; 50-1043/Product: desmoglein #status predicted <MAT>
F; 50-548/Domain: extracellular #status predicted <EXT>
F; 52-157/Domain: cadherin repeat homology <CR1>
F; 160-269/Domain: cadherin repeat homology <CR2>
F; 272-385/Domain: cadherin repeat homology <CR3>
F; 392-491/Domain: cadherin repeat homology <CR4>
F; 548-574/Domain: transmembrane #status predicted <TMM>
F; 575-1043/Domain: intracellular #status predicted <INT>
F; 846-875/Domain: desmoglein repeat <DG1>
F; 876-905/Domain: desmoglein repeat <DG2>
F; 906-933/Domain: desmoglein repeat <DG3>
F; 934-962/Domain: desmoglein repeat <DG4>
F; 963-1012/Region: glycine/serine-rich
F; 110/Binding site: carbohydrate (Asn) (covalent) #status experimental
F; 180,496/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 83.5%; Score 81; DB 1; Length 1043;
Best Local Similarity 78.9%; Pred. No. 1.2e-05;
Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EPNHNSKIAPKIVSQEPA 19
|||||:|||||:|||||:
Db 186 EPNHNSKIAPKIVSQEPA 204

RESULT 3
10HUG1

desmoglein 1 precursor - human
N; Alternate names: desmosomal glycoprotein 1
C; Species: Homo sapiens (man)
C; Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 22-Jun-1999
C; Accession: S16906; A39706; A61254; A61279; S16158
R; Buxton, R.S.
submitted to the EMBL Data Library, November 1990
A; Reference number: S16906
A; Accession: S16906
A; Molecule type: mRNA
A; Residues: 1-1049 <BUX>
A; Cross-references: EMBL:X56654; NID:G30505; PIDN:CAA39976.1; PID:G30506
R; Wheeler, G.N.; Parker, A.E.; Thomas, C.L.; Atlantiotis, P.; Poynter, D.; Arnemann, J.;
Proc. Natl. Acad. Sci. U.S.A. 88, 4796-4800, 1991
A; Title: Desmosomal glycoprotein DGI, a component of intercellular desmosome junctions,
A; Reference number: A39706; MUID:91271279; PMID:1711210
A; Accession: A39706
A; Molecule type: mRNA
A; Residues: 24-1049 <WHE>
A; Cross-references: GB:X56654
R; Nilles, L.A.; Parry, D.A.D.; Powers, E.E.; Angst, B.D.; Wagner, R.M.; Green, K.J.
J. Cell Biol. 99, 809-821, 1991
A; Title: Structural analysis and expression of human desmoglein: a cadherin-like component of the desmosome
A; Reference number: A61254; MUID:92121251; PMID:1770008
A; Accession: A61254
A; Molecule type: mRNA
A; Residues: 26-1049 <NIL>
R; Wheeler, G.N.; Buxton, R.S.; Parker, A.E.; Arnemann, J.; Rees, D.A.; King, I.A.; Magee
Biochem. Soc. Trans. 19, 1060-1064, 1991
A; Title: Desmosomal glycoproteins I, II and III: novel members of the cadherin superfamily
A; Reference number: A61279; MUID:92175187; PMID:1794462
A; Accession: A61279
A; Status: not compared with conceptual translation
A; Molecule type: mRNA
A; Residues: 1-55 <WH3>
C; Genetics:
A; Gene: GDB:DSG1
A; Cross-references: GDB:126563; OMIM:125670
A; Map position: 18q12.1-18q12.2
C; Superfamily: cadherin; cadherin repeat homology
C; Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane protein
F; 1-23/Domain: signal sequence #status predicted <SIG>
F; 24-49/Domain: propeptide #status predicted <PRO>
F; 50-1049/Product: desmoglein #status predicted <MAT>
F; 50-548/Domain: extracellular #status predicted <EXT>
F; 52-157/Domain: cadherin repeat homology <CR1>
F; 160-269/Domain: cadherin repeat homology <CR2>
F; 272-385/Domain: cadherin repeat homology <CR3>
F; 392-493/Domain: cadherin repeat homology <CR4>
F; 509-530/Region: serine/threonine-rich
F; 549-569/Domain: transmembrane #status predicted <TMM>
F; 572-1049/Domain: intracellular #status predicted <INT>
F; 840-869/Domain: desmoglein repeat <DG1>
F; 870-899/Domain: desmoglein repeat <DG2>
F; 900-927/Domain: desmoglein repeat <DG3>
F; 928-956/Domain: desmoglein repeat <DG4>
F; 969-1019/Region: glycine/serine-rich
F; 110/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 83.5%; Score 81; DB 1; Length 1049;
Best Local Similarity 78.9%; Pred. No. 1.2e-05;
Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EPNHNSKIAPKIVSQEPA 19
|||||:|||||:|||||:
Db 186 EPNHNSKIAPKIVSQEPA 204

RESULT 4
S38673
desmoglein 2 - human
N; Alternate names: desmoglein HDGC
C; Species: Homo sapiens (man)

C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 20-Aug-1999
 A;Accession: S38673; B38872
 R;Zimbelmann, R.
 submitted to the EMBL Data Library, September 1993
 A;Reference number: S38673
 A;Accession: S38673
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-1117 <ZIM>
 A;Cross-references: EMBL:226317; NID:g416177; PIDN:CAA81226.1; PID:g416178
 R;Koch, P.J.; Goldschmidt, M.D.; Walsh, M.J.; Zimbelmann, R.; Franke, W.W.
 Eur. J. Cell Biol. 55, 200-208, 1991
 A;Title: Complete amino acid sequence of the epidermal desmoglein precursor polypeptide
 A;Reference number: A38872; MUID:92037656; PMID:1935985
 A;Accession: B38872
 A;Molecule type: mRNA
 A;Residues: 777-1117 <KOC>
 A;Cross-references: GB:S64273
 C;Genetics:
 A;Gene: GDB:DSG2
 A;Cross-references: GDB:128808; OMIM:125671
 A;Map position: 18q12.1-18q12.2
 C;Superfamily: cadherin; cadherin repeat homology
 C;Keywords: calcium binding; cell adhesion; duplication; glycoprotein; membrane protein
 F;51-158/Domain: cadherin repeat homology <CR1>
 F;161-271/Domain: cadherin repeat homology <CR2>

Query Match 73.2%; Score 71; DB 2; Length 1117;
 Best Local Similarity 73.7%; Pred. No. 0.00072;
 Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EPNHLNKAIFKIVSQEPA 19
 ||| |||||:::|||||
 Db 187 EPNTLNKISVRIIVSLRPA 205

RESULT 5
 S57904
 virR49 protein - Streptococcus pyogenes (strain CS101, serotype M49)
 C;Species: Streptococcus pyogenes
 A;Variety: Strain CS101, serotype M49
 C;Date: 29-Nov-1995 #sequence_revision 23-Feb-1996 #text_change 20-Jun-2000
 A;Accession: S57904; S34297; S60059
 R;Podbielski, A.; Flosdorff, A.; Weber-Heynenann, J.
 Infect. Immun. 63, 9-20, 1995
 A;Title: The group A streptococcal virR49 gene controls expression of four structural vi
 A;Reference number: S57904; MUID:95105032; PMID:7806389
 A;Accession: S57904
 A;Status: nucleic acid sequence not shown
 A;Molecule type: DNA
 A;Residues: 1-533 <POD>
 A;Cross-references: EMBL:X68501
 R;Podbielski, A.; Weber-Heynenann, J.
 submitted to the EMBL Data Library, September 1992
 A;Description: VirR gene of DF-positive group A streptococci.
 A;Reference number: S34297
 A;Accession: S34297
 A;Molecule type: DNA
 A;Residues: 1-458, 'V', 460-483, 'KIKRD', 489-533 <POW>
 A;Cross-references: EMBL:X68501
 R;Podbielski, A.
 submitted to the EMBL Data Library, February 1995
 A;Reference number: S60059
 A;Accession: S60059
 A;Molecule type: DNA
 A;Residues: 1-238, 'H', 240-329, 'D', 331-533 <POF>
 A;Cross-references: EMBL:X68501; NID:g677864; PIDN:CAA48513.1; PID:g677866
 C;Genetics:
 A;Gene: virR49
 C;Superfamily: Streptococcus positive regulatory protein Mga

Query Match 53.6%; Score 52; DB 2; Length 533;
 Best Local Similarity 50.0%; Pred. No. 0.65;
 Matches 10; Conservative 5; Mismatches 1; Indels 4; Gaps 1;

Qy 3 NHLNSKI----AFKIVSQEP 18
 |||||:|:|:|:|:|:|
 Db 269 NHLNDKLEIGCAFEINQDP 288

RESULT 6
 JC5040
 positive regulatory protein Mga - Streptococcus pyogenes
 C;Species: Streptococcus pyogenes
 C;Date: 21-Jan-1997 #sequence_revision 21-Jan-1997 #text_change 01-Dec-2000
 A;Accession: JC5040; S52535
 R;Andersson, G.; McIver, K.; Heden, L.O.; Scott, J.R.
 Gene 175, 77-81, 1996
 A;Title: Complementatation of divergent mga genes in group A Streptococcus.
 A;Reference number: JC5040; MUID:97074652; PMID:8917079
 A;Accession: JC5040
 A;Status: nucleic acid sequence not shown
 A;Molecule type: DNA
 A;Residues: 1-533 <AND>
 A;Cross-references: EMBL:X95408; NID:gl246851; PIDN:CAA64690.1; PID:gl246852
 A;Experimental source: strain AP4
 R;Katerov, V.; Schalen, C.; Totolian, A.A.
 Mol. Gen. Genet. 245, 78-85, 1994
 A;Title: Sequencing of genes within the vir regulon of Streptococcus pyogenes type M15
 A;Reference number: S52535; MUID:95147851; PMID:7845360
 A;Accession: S52535
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 299-360, 'P', 362-533 <KAT>
 A;Cross-references: GB:S75411; NID:g914107; PIDN:AAB33260.1; PID:g914108
 C;Comment: This protein acts as the second component of a signal transducing system.
 C;Genetics:
 A;Gene: mga4
 C;Superfamily: Streptococcus positive regulatory protein Mga

Query Match 53.6%; Score 52; DB 2; Length 533;
 Best Local Similarity 50.0%; Pred. No. 0.65;
 Matches 10; Conservative 5; Mismatches 1; Indels 4; Gaps 1;

Qy 3 NHLNSKI----AFKIVSQEP 18
 |||||:|:|:|:|:|:|
 Db 269 NHLNDKLEIGCAFEINQDP 288

RESULT 7
 IJHUCN
 cadherin 2 precursor - human
 N;Alternate names: N-cadherin; neuronal cadherin
 C;Species: Homo sapiens (man)
 C;Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 22-Jun-1999
 A;Accession: A38870; S11487; JQ0751; S13799
 R;Reid, R.A.
 submitted to the EMBL Data Library, November 1990
 A;Reference number: A38870
 A;Accession: A38870
 A;Molecule type: mRNA
 A;Residues: 1-906 <REI>
 A;Cross-references: EMBL:X54315; NID:g34998; PIDN:CAA38213.1; PID:g34999
 R;Reid, R.A.; Hemperly, J.J.
 Nucleic Acids Res. 18, 5896, 1990
 A;Title: Human N-cadherin: nucleotide and deduced amino acid sequence.
 A;Reference number: S11487; MUID:91016946; PMID:2216790
 A;Accession: S11487
 A;Molecule type: mRNA
 A;Residues: 1-340, 'N', 342-698, 'R', 700-704, 'P', 706-906 <RE2>
 A;Cross-references: EMBL:X54315
 A;Note: this sequence has been revised in reference A38870
 R;Walsh, F.S.; Barton, C.H.; Putt, W.; Moore, S.E.; Kelsell, D.; Spurr, N.; Goodfellow, J.
 J. Neurochem. 55, 805-812, 1990
 A;Title: N-cadherin gene maps to human chromosome 18 and is not linked to the E-cadherin
 A;Reference number: JQ0751; MUID:90347462; PMID:2384753

```
F;865-878/Region: serine-rich
F;190,273,325,402,572,651,692/Binding site: carbohydrate (Asn) (covalent) #status pred
```

	Query March	50.5%	Score 49;	DB 1;	Length 906;
Best Local Similarity	42.1%	Pred. No. 3.9;			
Matches	8;	Conservative	6;	Mismatches	5;
				Indels	0;
Gaps					0;

Qy 1 EPNHLNSKIAPKIVSQEPA 19
::|||::|::|::|::|:

Db 296 DFNALNGMLRYRIISQAQS 314
::|::|::|::|::|::|:

RESULT 10
I46536
Ksp-cadherin - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 05-Nov-1999
C:Accession: I46536
J.Thomson, R.B.; Igarashi, P.; Bienesderfer, D.; Kim, R.; Abu-Alfa, A.; Soleimani, M.;
J. Biol. Chem. 270, 17594-17601, 1995
A>Title: Isolation and cDNA cloning of Ksp-cadherin, a novel kidney-specific member of
A:Reference number: I46536; PMID:95340560; EMBL:7615566
A:Accession: I46536
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-829 <THO>
A:Cross-references: EMBL:I208945; NID:g902885; PIDN:AAC48472.1; PID:g902886

	Query Match	49.5%	Score 48;	DB 2;	Length 829;
Best Local Similarity	52.6%	Pred. No. 5.3;			
Matches	11;	Conservative	2;	Mismatches	7;
				Indels	0;
Gaps					0;

```

QY      1 EPNHLSKIAFKIVSQEPA 19
      ||| : ||| : ||| : |||
Db      155 EFGTANSLRFLHSQTQA 173

RESULT 11
B96963
uncharacterized conserved protein, similar to B. subtilis yvcL CAC0513 [imported] - Clos
C/Species: Clostridium acetobutylicum
C/Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
R/Accession: B96963
R/Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A/Reference number: A96900; MUID:21359325; PMID:21359325
A/Accession: B96963
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-317 <KUR>
A/Cross-references: GB:AE001437; PIDN:AAK78493.1; PID:gi5023376; GSPDB:GN00168
A/Experimental source: Clostridium acetobutylicum ATCC824
C/Genetics:
A/Gene: CAC0513
C/Superfamily: Streptomyces coelicolor hypothetical protein SCC54.10c

Query Match      48.5%; Score 47; DB 2; Length 317;
Best Local Similarity 57.1%; Pred. No. 2.7;
Matches      8; Conservative      4; Mismatches      2; Indels      0; Gaps      0;

QY      6 NSKIAFKIVSQEPA 19
      ||| : ||| : ||| : |||
Db      41 NRKLSFKIITENPA 54

RESULT 12
I50116
N-cadherin precursor - zebra fish
C/Species: Brachydanio rerio (zebra fish)
C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jan-2000
C/Accession: I50116
Dev. Dyn. 201, 121-136, 1994.
A/Title: Structure and distribution of N-cadherin in developing zebrafish embryos: morph
A/Reference number: I50116; MUID:95178741; PMID:7873785
A/Accession: I50116
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: mRNA
A/Residues: 1-783 <BIT>
A/Cross-references: EMBL:X67648; NID:g623209; PIDN:CAA47890.1; PID:g623210
C/Superfamily: cadherin; cadherin repeat homology
F:146-258/Domain: cadherin repeat homology <CDH>

Query Match      47.4%; Score 46; DB 2; Length 783;
Best Local Similarity 44.4%; Pred. No. 11;
Matches      8; Conservative      4; Mismatches      6; Indels      0; Gaps      0;

QY      1 EPNHLSKIAFKIVSQEP 18
      ||| : ||| : ||| : |||
Db      172 DPNTANGMLRYKILSQTP 189

RESULT 13
A53731
translation initiation factor eIF-2 alpha chain kinase (EC 2.7.1.1) - rat
N/Alternate names: heme-controlled repressor
C/Species: Rattus norvegicus (Norway rat)
C/Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 24-Sep-1999
C/Accession: A53731
R/Mellor, H.; Flowers, K.M.; Kimball, S.R.; Jefferson, L.S.
J. Biol. Chem. 269, 10201-10204, 1994
A/Title: Cloning and characterization of cDNA encoding rat hemin-sensitive initiation fa
A/Reference number: A53731; MUID:94193700; PMID:7908290

```

```

A/Accession: A53731
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-620 <MEL>
A/Cross-references: GB:I27707; NID:g443688; PIDN:AAA18255.1; PID:g443689
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homo
C/Keywords: phosphotransferase
F:165-581/Domain: protein kinase homology <KIN>

Query Match      46.4%; Score 45; DB 2; Length 620;
Best Local Similarity 56.2%; Pred. No. 13;
Matches      9; Conservative      3; Mismatches      4; Indels      0; Gaps      0;

QY      1 EPNHLSKIAFKIVSQ 16
      ||| : ||| : ||| : |||
Db      81 EFNPLHSKQVFKLLCQ 96

RESULT 14
LJMSCE
E-cadherin precursor, epithelial - mouse
N/Alternate names: uvomorulin
C/Species: Mus musculus (house mouse)
C/Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 22-Jun-1999
C/Accession: S04528; S03160; I49565; S48735
R/Nagafuchi, A.; Shirayoshi, Y.; Okazaki, K.; Yasuda, K.; Takeichi, M.
Nature 329, 341-343, 1987
A/Title: Transformation of cell adhesion properties by exogenously introduced E-cadheri
A/Reference number: S04528; MUID:87315445; PMID:3498123
A/Accession: S04528
A/Molecule type: mRNA
A/Residues: 1-412, 'V', 414-884 <NAG>
A/Cross-references: EMBL:X06115
R/Ringwald, M.; Schuh, R.; Vestweber, D.; Eistetter, H.; Lottspeich, F.; Engel, J.; Doe
EMBO J. 6, 3647-3653, 1987
A/Title: The structure of cell adhesion molecule uvomorulin. Insights into the molecula
A/Reference number: S03160; MUID:88111553; PMID:3501370
A/Accession: S03160
A/Molecule type: mRNA
A/Residues: 157-884 <RIN>
A/Cross-references: EMBL:X06339
A/Note: part of this sequence, including the amino end of the mature protein, was confi
R/Behrens, J.; Loewrick, O.; Klein-Hitpass, L.; Birchmeier, W.
Proc. Natl. Acad. Sci. U.S.A. 88, 11495-11499, 1991
A/Title: The E-cadherin promoter: Functional analysis of a G-C-rich region and an epith
A/Reference number: I49565; MUID:92107977; PMID:1763063
A/Accession: I49565
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-15 <RES>
A/Cross-references: GB:M81449; NID:gi92325; PIDN:AAA37352.1; PID:gi92326
R/Tong, K.I.; Yau, P.; Overduin, M.; Bagby, S.; Porumb, T.; Takeichi, M.; Ikura, M.
FEBS Lett. 352, 318-322, 1994
A/Title: Purification and spectroscopic characterization of a recombinant amino-termina
A/Reference number: S48735; MUID:95010732; PMID:7925993
A/Accession: S48735
A/Status: preliminary
A/Molecule type: protein
A/Residues: 156-300 <TON>
C/Comment: Cadherins mediate calcium-dependent intercellular adhesion, and are thought
C/Genetics:
A/Gene: E-cadherin
C/Superfamily: cadherin; cadherin repeat homology
C/Keywords: calcium binding; cell adhesion; duplication; glycoprotein; transmembrane pro
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-156/Domain: propeptide #status predicted <PRO>
F:157-884/Product: E-cadherin, epithelial #status experimental <MAT>
F:157-699/Domain: extracellular #status predicted <EXT>
F:159-264/Domain: cadherin repeat homology <CR1>
F:234-239/Region: cadherin binding #status predicted
F:267-377/Domain: cadherin repeat homology <CR2>
F:380-488/Domain: cadherin repeat homology <CR3>
F:489-597/Domain: cadherin repeat homology <CR4>

```

Search completed: October 12, 2004, 05:59:01
Job time : 28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2004, 05:56:44 ; Search time 15 Seconds
(without alignments)
65.956 Million cell updates/sec

Title: US-10-799-005A-1

Perfect score: 97

Sequence: 1 EPNHLSKIAFKIVSQEPA 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	97	100.0	999	1	DSG3_HUMAN
2	95	97.9	993	1	DSG3_MOUSE
3	81	83.5	1043	1	DSG1_BOVIN
4	81	83.5	1049	1	DSG1_HUMAN
5	71	73.2	1117	1	DSG2_HUMAN
6	65	67.0	1122	1	DSG2_MOUSE
7	52	53.6	906	1	CAD2_HUMAN
8	49	50.5	238	1	CAD2_CRIGR
9	49	50.5	877	1	CAD2_MOUSE
10	49	50.5	906	1	CAD2_BOVIN
11	49	50.5	906	1	CAD2_MOUSE
12	48	49.5	829	1	CAD2_RAT
13	48	49.5	829	1	CADG_HUMAN
14	46	47.4	239	1	HIS1_PASMO
15	46	47.4	827	1	CADH_MOUSE
16	46	47.4	893	1	CAD2_BRARE
17	45	46.4	459	1	DNAI_NITEU
18	45	46.4	619	1	E2K1_MOUSE
19	45	46.4	620	1	E2K1_RAT
20	45	46.4	884	1	CAD1_MOUSE
21	45	46.4	1321	1	MDRI_CABEL
22	44	45.4	905	1	CADN_XENLA
23	44	45.4	906	1	CADO_XENLA
24	43	44.3	145	1	CAD3_PIG
25	43	44.3	307	1	DIM1_SCHPO
26	43	44.3	725	1	RNR_MYCPO
27	43	44.3	827	1	CADH_RAT
28	43	44.3	863	1	DSC2_BOVIN
29	43	44.3	912	1	CAD2_CHICK
30	42.5	43.8	181	1	CAV1_FUGRU
31	42	43.3	479	1	PAP1_VACCC
32	42	43.3	479	1	PAP1_VACCV
33	42	43.3	479	1	PAP1_VARV

34 42 43.3 882 1 CAD1_HUMAN P12830 homo sapien
35 42 43.3 886 1 CAD1_RAT Q970t4 rattus norv
36 42 43.3 893 1 DSC1_BOVIN Q01107 bos taurus
37 42 43.3 894 1 DSC1_HUMAN Q08554 homo sapien
38 41 42.3 626 1 E2K1_RABIT P31279 o eukaryoti
39 41 42.3 630 1 E2K1_HUMAN Q9pqi3 h eukaryoti
40 41 42.3 732 1 CADL_CHICK P33145 gallus gall
41 41 42.3 886 1 DSC1_MOUSE P55849 mus musculu
42 40 41.2 102 1 YBH1_YEAST P38185 saccharomyc
43 40 41.2 228 1 CLCB_BOVIN P04975 bos taurus
44 40 41.2 229 1 CLCB_HUMAN P09475 homo sapien
45 40 41.2 229 1 CLCB_RAT P08082 rattus norv

ALIGNMENTS

RESULT 1
DSG3_HUMAN
ID DSG3_HUMAN STANDARD; PRT; 999 AA.
AC P32926;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Desmoglein 3 precursor (130 kDa pemphigus vulgaris antigen) (PVA).
GN DSG3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92069753; PubMed=1720352;
RA Amagai M., Klaus-Kovtun V., Stanley J.R.;
RT "Autoantibodies against a novel epithelial cadherin in pemphigus
vulgaris, a disease of cell adhesion.";
RL Cell 67:869-877(1991).
CC -!- FUNCTION: Component of intercellular desmosome junctions. Involved
in the interaction of plaque proteins and intermediate filaments
mediating cell-cell adhesion.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: Epidermis, tongue, tonsil, esophagus and
carcinomas.
CC -!- DOMAIN: Calcium may be bound by the cadherin-like repeats
(Potential).
CC -!- DISEASE: Pemphigus vulgaris (PV) is a potentially lethal skin
disease in which epidermal blisters occur as the result of the
loss of cell-cell adhesion caused by the action of autoantibodies
against desmoglein 3.
CC -!- SIMILARITY: Contains 4 cadherin domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; M76482; AAA60230.1; -;
DR PIR; A41088; IJHUG3.
DR HSSP; P15116; INCU.
DR Genew; HGNC:3050; DSG3.
DR MIM; 169615; -;
DR InterPro; IPR002126; Cadherin.
DR Pfam; PF00028; cadherin; 4.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 4.
DR PROSITE; PS00232; CADHERIN 1; 3.
DR PROSITE; PS00268; CADHERIN 2; 4.
KW Cell adhesion; Signal; Transmembrane; Cytoskeleton; Glycoprotein;
KW Calcium-binding; Repeat.
FT SIGNAL 1 23 POTENTIAL.

DR PROSITE; PS0268; CADHERIN 2; 4.
KW Cell adhesion; Signal; Transmembrane; Cytoskeleton; Glycoprotein;
KW Calcium-binding; Repeat.
FT SIGNAL 1 23 POTENTIAL.
FT PROPEP 24 49
FT DOMAIN 50 615
FT TRANSMEM 616 640
FT DOMAIN 641 999
FT CYTOPLASMIC (POTENTIAL).
FT DOMAIN 159 268
FT CADHERIN 1.
FT DOMAIN 269 383
FT CADHERIN 2.
FT DOMAIN 386 499
FT CADHERIN 3.
FT REPEAT 910 935
FT DESMOGLEIN REPEAT 1.
FT REPEAT 936 966
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 110 180
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 180 459
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 459 545
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 545 545
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 999 AA; 107503 MW; 60479DD46AC219A1 CRC64;

Query Match 100.0%; Score 97; DB 1; Length 999;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EPNHLSKIAPKIVSQEPA 19
Db 186 EPNHLSKIAPKIVSQEPA 204

RESULT 2
DSG3 MOUSE
ID DSG3 MOUSE STANDARD; PRT; 993 AA.
AC O35902;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Desmoglein 3 precursor (130 kDa pemphigus vulgaris antigen homolog)
DE (Fragment).
GN DSG3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/c;
RX MEDLINE=95201440; PubMed=7894164;
RA Ishikawa H., Silos S.A., Tamai K., Copeland N.G., Gilbert D.J.,
Jenkins N.A., Uitto J.;
RT "CDNA cloning and chromosomal assignment of the mouse gene for
desmoglein 3 (Dsg3), the pemphigus vulgaris antigen.";
RL Mamm. Genome 5:803-804(1994).
CC
CC -!- FUNCTION: Component of intercellular desmosome junctions. Involved
in the interaction of plaque proteins and intermediate filaments
mediating cell-cell adhesion.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- DOMAIN: Calcium may be bound by the cadherin-like repeats
(Potential).
CC -!- SIMILARITY: Contains 4 cadherin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; U86016; AAB65091.1; -.
CC HSSP; P15116; 1NCJ.
CC MGD; MGI:99499; Dsg3.
CC InterPro; IPR002126; Cadherin.
CC Pfam; PF00028; cadherin; 4.
CC PRINTS; PR00205; CADHERIN.
CC SMART; SM00112; CA; 4.
CC PROSITE; PS00232; CADHERIN_1; 2.

DR PROSITE; PS0268; CADHERIN 2; 4.
KW Cell adhesion; Signal; Transmembrane; Cytoskeleton; Glycoprotein;
KW Calcium-binding; Repeat.
FT SIGNAL 1 23 POTENTIAL.
FT PROPEP 24 49
FT DOMAIN 50 615
FT TRANSMEM 616 640
FT CYTOPLASMIC (POTENTIAL).
FT DOMAIN 159 268
FT CADHERIN 1.
FT DOMAIN 269 383
FT CADHERIN 2.
FT DOMAIN 386 499
FT CADHERIN 3.
FT REPEAT 910 935
FT DESMOGLEIN REPEAT 1.
FT REPEAT 936 966
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 110 180
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 180 459
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 459 545
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 545 545
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 999 AA; 107503 MW; 60479DD46AC219A1 CRC64;

Query Match 100.0%; Score 97; DB 1; Length 999;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EPNHLSKIAPKIVSQEPA 19
Db 186 EPNHLSKIAPKIVSQEPA 204

RESULT 3
DSG1 BOVIN
ID DSG1 BOVIN STANDARD; PRT; 1043 AA.
AC Q03763; Q9TS15;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Desmoglein 1 precursor (Desmosomal glycoprotein 1) (DGI) (DGI)
DE (Pemphigus foliaceus antigen).
GN DSG1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muzzle epithelium;
RC Koch P.J., Goldschmidt M.D., Zimbelmann R., Franke W.W.;
RA Submitted (MAR-1991) to the EMBL/GenBank/DBJ databases.
RL
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Thymus;
RX MEDLINE=94124543; PubMed=8294446;
RA Puttagunta S., Mathur M., Cowin P.;
RT "Structure of DSG1, the bovine desmosomal cadherin gene encoding the
pemphigus foliaceus antigen. Evidence of polymorphism.";
RL J. Biol. Chem. 269:1949-1955(1994).
RN [3]
RP SEQUENCE OF 44-1043 FROM N.A.
RC TISSUE=Muzzle epithelium;
RX MEDLINE=91168965; PubMed=1706270;
RA Koch P.J., Walsh M.J., Schmelz M., Goldschmidt M.D.,
Zimbelmann R., Franke W.W.;
RT "Identification of desmoglein, a constitutive desmosomal
glycoprotein, as a member of the cadherin family of cell adhesion
molecules.";
RL Eur. J. Cell Biol. 53:1-12(1990).
RN [4]
RP REVISIONS, AND SEQUENCE OF 101-123.
RX MEDLINE=92037656; PubMed=1935985;
RA Koch P.J., Goldschmidt M.D., Walsh M.J., Zimbelmann R.,

RA Franke W.W.;
 RT "Complete amino acid sequence of the epidermal desmoglein precursor
 RL polypeptide and identification of a second type of desmoglein gene."; Eur. J. Cell Biol. 55:200-208(1991).
 RN [5].
 RP SEQUENCE OF 44-493 FROM N.A.
 RX MEDLINE=91097553; PubMed=1702628;
 RA Goodwin L., Hill J.E., Raynor K., Raszi L., Manabe M., Cowin P.;
 RT "Desmoglein shows extensive homology to the cadherin family of cell
 RL adhesion molecules";
 RL Biochem. Biophys. Res. Commun. 173:1224-1230(1990).
 CC -!- FUNCTION: Component of intercellular desmosome junctions. Involved
 CC in the interaction of plaque proteins and intermediate filaments
 CC mediating cell-cell adhesion.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: Epidermis, muzzle, tongue and esophagus.
 CC -!- DOMAIN: Calcium may be bound by the cadherin-like repeats
 CC (Potential).
 CC -!- SIMILARITY: Contains 5 cadherin domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X58466; CAA41380.1; -;
 DR EMBL; X57784; CAA40930.1; -;
 DR EMBL; M58165; AAA62709.1; -;
 DR PIR; S14603; IJBOGI.
 DR HSSP; P15116; INCUJ.
 DR InterPro; IPR002126; Cadherin.
 DR InterPro; IPR002233; Cadherin_C_term.
 DR Pfam; PF00028; cadherin; 3.
 DR Pfam; PF01049; Cadherin_C_term; 1.
 DR SMART; SM00112; CA; 3.
 DR PROSITE; PS00232; CADHERIN_1; 2.
 DR PROSITE; PS0266; CADHERIN_2; 4.
 KW Cell adhesion; Signal; Transmembrane; Cytoskeleton; Glycoprotein;
 KW Calcium-binding; Repeat.
 FT SIGNAL 1 23 POTENTIAL.
 FT PROPEP 24 49 POTENTIAL.
 FT CHAIN 50 1043 DESMOGLEIN 1.
 FT DOMAIN 50 548 EXTRACELLULAR (POTENTIAL).
 FT TRANSEM 549 573 POTENTIAL.
 FT DOMAIN 574 1043 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 50 158 CADHERIN 1.
 FT DOMAIN 159 270 CADHERIN 2.
 FT DOMAIN 271 385 CADHERIN 3.
 FT DOMAIN 386 498 CADHERIN 4.
 FT REPEAT 819 845 DESMOGLEIN REPEAT 1.
 FT REPEAT 846 875 DESMOGLEIN REPEAT 2.
 FT REPEAT 876 905 DESMOGLEIN REPEAT 3.
 FT REPEAT 906 933 DESMOGLEIN REPEAT 4.
 FT REPEAT 934 962 DESMOGLEIN REPEAT 5.
 FT DOMAIN 963 1012 GLY/SER-RICH.
 FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 496 496 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 124 124 I -> V (IN REF. 2 AND 5).
 FT CONFLICT 889 889 M -> N (IN REF. 2).
 FT CONFLICT 945 945 M -> N (IN REF. 2).
 SQ SEQUENCE 1043 AA; 112243 MW; ADE46133F8B77C11 CRC64;
 Query Match 83.5%; Score 81; DB 1; Length 1043;
 Best Local Similarity 78.9%; Pred. No. 7.9e-06;
 Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 EPNHLSKIAFKIVSQEPA 19
 ||||:|||||:||||: |||:

Db 186 EPNHLSKIAFKIIRQEPS 204
 RESULT 4
 ID DSG1_HUMAN STANDARD; PRT; 1049 AA.
 AC Q02413;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Desmoglein 1 precursor (Desmosomal glycoprotein 1) (DGI) (DGI)
 GN DSG1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Keratinocytes;
 RX MEDLINE=91271279; PubMed=1711210;
 RA Wheeler G.N., Parker A.E., Thomas C.L., Ataliotis P., Poynter D.,
 RA Arneemann J., Rutman A.J., Pidsley S.C., Watt F.M., Rees D.A.,
 RA Buxton R.S., Magee A.I.;
 RT "Desmosomal glycoprotein DGI, a component of intercellular desmosome
 RT junctions, is related to the cadherin family of cell adhesion
 RT molecules."; Proc. Natl. Acad. Sci. U.S.A. 88:4796-4800(1991).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Foreskin;
 RC MEDLINE=92121251; PubMed=1770008;
 RA Nilles L.A., Parry D.A., Powers E.E., Angst B.D., Wagner R.M.,
 RA Green K.J.;
 RT "Structural analysis and expression of human desmoglein: a cadherin-
 RT like component of the desmosome."; J. Cell Sci. 99:809-821(1991).
 RL [1]
 CC -!- FUNCTION: Component of intercellular desmosome junctions. Involved
 CC in the interaction of plaque proteins and intermediate filaments
 CC mediating cell-cell adhesion.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: Epidermis, tongue, tonsil and esophagus.
 CC -!- DOMAIN: Calcium may be bound by the cadherin-like repeats
 CC (Potential).
 CC -!- SIMILARITY: Contains 5 cadherin domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X56654; CAA39976.1; -;
 DR EMBL; AF097935; AAC83817.1; -;
 DR PIR; S16906; IJHUGL.
 DR HSSP; P15116; INCUJ.
 DR Genew; HGNC:3048; DSG1.
 DR MIM; 125670; -;
 DR GO; GO:0007155; P:cell adhesion; TAS.
 DR InterPro; IPR002126; Cadherin.
 DR InterPro; IPR000233; Cadherin_C_term.
 DR Pfam; PF00028; cadherin; 4.
 DR Pfam; PF01049; Cadherin_C_term; 1.
 DR PRINTS; PR00205; CADHERIN_C.
 DR SMART; SM00112; CA; 4.
 DR PROSITE; PS00232; CADHERIN_1; 2.
 DR PROSITE; PS0266; CADHERIN_2; 4.
 KW Cell adhesion; Signal; Transmembrane; Cytoskeleton; Glycoprotein;
 KW Calcium-binding; Repeat.
 FT SIGNAL 1 23 POTENTIAL.
 FT PROPEP 24 49 POTENTIAL.

FT CHAIN 50 1049 DESMOGLEIN 1.
 FT DOMAIN 50 545 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 546 570 POTENTIAL.
 FT DOMAIN 571 1049 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 50 158 CADHERIN 1.
 FT DOMAIN 159 270 CADHERIN 2.
 FT DOMAIN 271 385 CADHERIN 3.
 FT DOMAIN 386 497 CADHERIN 4.
 FT DOMAIN 813 839 DESMOGLEIN REPEAT 1.
 FT REPEAT 840 869 DESMOGLEIN REPEAT 2.
 FT REPEAT 870 899 DESMOGLEIN REPEAT 3.
 FT REPEAT 900 927 DESMOGLEIN REPEAT 4.
 FT REPEAT 928 956 DESMOGLEIN REPEAT 5.
 FT DOMAIN 969 1019 GLY/SER-RICH.
 FT CARBOHYD 36 36 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 110 110 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 180 180 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 1049 AA; 113715 MW; BEE125655B9D6619 CRC64;

Query Match 83.5%; Score 81; DB 1; Length 1049;
 Best Local Similarity 78.9%; Pred. No. 7.9e-06;
 Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFKIVSQEPA 19
 |||:|||||||: |||:
 Db 186 EPNHLSKIAFKIIQEPS 204

RESULT 5
 DSG2 HUMAN
 ID DSG2 HUMAN STANDARD; PRT; 1117 AA.
 AC Q14126;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Desmoglein 2 precursor (HDGC).
 GN DSG2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Colon carcinoma;
 RX MEDLINE=94192736; PubMed=8143788;
 RA Schaefer S., Koch P.J., Franke W.W.;
 RT "Identification of the ubiquitous human desmoglein, Dsg2, and the
 RT expression catalogue of the desmoglein subfamily of desmosomal
 RT cadherins.";
 RL Exp. Cell Res. 211:391-399(1994).
 RN [2]
 RP SEQUENCE OF 777-1117 FROM N.A.
 RX MEDLINE=92037856; PubMed=1935985;
 RA Koch P.J., Goldschmidt M.D., Walsh M.J., Zimbelmann R., Franke W.W.;
 RT "Complete amino acid sequence of the epidermal desmoglein precursor
 RT polypeptide and identification of a second type of desmoglein gene.";
 RL Eur. J. Cell Biol. 55:200-208(1991).
 RN [3]
 RP CARBOHYDRATE-LINKAGE SITE ASN-111.
 RX MEDLINE=22660472; PubMed=12754519;
 RA Zhang H., Li X.-J., Martin D.B., Aebersold R.;
 RT "Identification and quantification of N-linked glycoproteins using
 RT hydrazide chemistry, stable isotope labeling and mass spectrometry.";
 RL Nat. Biotechnol. 21:660-666(2003).
 CC -!- FUNCTION: Component of intercellular desmosome junctions. Involved
 CC in the interaction of plaque proteins and intermediate filaments
 CC mediating cell-cell adhesion.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
 CC -!- TISSUE SPECIFICITY: All of the tissues tested and carcinomas.
 CC -!- DOMAIN: Calcium may be bound by the cadherin-like repeats
 CC (Potential).
 CC -!- SIMILARITY: Contains 4 cadherin domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z26317; CAA81226.1; -
 DR PIR; S38673; S38673.
 DR HSSP; F15116; INCI.
 DR Genew; HGNC:3049; DSG2.
 DR MIM; 125671; -
 DR GO; GO:0005911; C:intercellular junction; TAS.
 DR InterPro; IPR002126; Cadherin.
 DR Pfam; PF00028; cadherin; 4.
 DR PRINTS; PR00205; CADHERIN.
 DR SMART; SM00112; CA; 4.
 DR PROSITE; PS00232; CADHERIN 1; 3.
 DR PROSITE; PS00268; CADHERIN 2; 4.
 KW Cell adhesion; Glycoprotein; Transmembrane; Repeat; Signal;
 KW Cytoskeleton; Calcium-binding.
 FT SIGNAL 1 23 POTENTIAL.
 FT PROPEP 24 48 POTENTIAL.
 FT CHAIN 49 1117 DESMOGLEIN 2.
 FT DOMAIN 49 608 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 609 633 POTENTIAL.
 FT DOMAIN 634 1117 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 49 159 CADHERIN 1.
 FT DOMAIN 160 272 CADHERIN 2.
 FT DOMAIN 273 387 CADHERIN 3.
 FT DOMAIN 388 502 CADHERIN 4.
 FT REPEAT 880 911 DESMOGLEIN REPEAT 1.
 FT REPEAT 912 941 DESMOGLEIN REPEAT 2.
 FT REPEAT 942 967 DESMOGLEIN REPEAT 3.
 FT REPEAT 968 991 DESMOGLEIN REPEAT 4.
 FT REPEAT 992 1020 DESMOGLEIN REPEAT 5.
 FT REPEAT 1021 1050 DESMOGLEIN REPEAT 6.
 FT CARBOHYD 111 111 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 181 181 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 308 308 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 461 461 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 513 513 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 1117 AA; 122385 MW; 223B897FED70B289 CRC64;

Query Match 73.2%; Score 71; DB 1; Length 1117;
 Best Local Similarity 73.7%; Pred. No. 0.00044;
 Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFKIVSQEPA 19
 |||:|||||||: |||:
 Db 187 EPNHLSKIVSVLEPA 205

RESULT 6
 DSG2 MOUSE
 ID DSG2 MOUSE STANDARD; PRT; 1122 AA.
 AC O5511; Q8K069; Q8R517;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Desmoglein 2 precursor.
 GN DSG2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RA Zhuxiang N., Garrard D.R.;
 RT "Desmosomal cadherins mediate homophilic cell adhesion.";
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.


```
CC manner in connecting cells; cadherins may thus contribute to the
CC sorting of heterogeneous cell types. N-cadherin may be involved in
CC neuronal recognition mechanism.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: Contains 5 cadherin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X57548; CAA40773.1; -.
DR EMBL; X54315; CAA38213.1; -.
DR EMBL; S42303; AAB22854.1; -.
DR EMBL; M34064; AAA03236.1; -.
DR EMBL; Z27420; CAA81799.1; -.
DR FIR; A38870; IJHUCN.
DR HSP; P15116; INCU.
DR Genew; HGNC:1759; CDH2.
DR MIM; 114020; -.
DR GO; GO:0007155; P:cell adhesion; TAS.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR000233; Cadherin_C_term.
DR Pfam; PF00028; cadherin; 5.
DR Pfam; PF01049; Cadherin_C_term; 1.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 5.
DR PROSITE; PS00232; CADHERIN_1; 3.
DR PROSITE; PS0268; CADHERIN_2; 5.
DR Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;
KW Signal.
FT SIGNAL 1 23 POTENTIAL.
FT PROPEP 24 159
FT CHAIN 160 906 NEURAL-CADHERIN.
FT DOMAIN 160 724 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 725 746 POTENTIAL.
FT DOMAIN 747 906 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 160 267 CADHERIN 1.
FT DOMAIN 268 382 CADHERIN 2.
FT DOMAIN 383 497 CADHERIN 3.
FT DOMAIN 498 603 CADHERIN 4.
FT DOMAIN 604 714 CADHERIN 5.
FT DOMAIN 863 878 SER-RICH.
FT CARBOHYD 190 190 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 273 273 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 325 325 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 572 572 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 651 651 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 692 692 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 12 12 MISSING (IN REF. 3).
FT CONFLICT 16 16 L -> A (IN REF. 3 AND 5).
FT CONFLICT 196 196 S -> T (IN REF. 1).
FT CONFLICT 212 212 I -> L (IN REF. 4).
FT CONFLICT 357 357 N -> I (IN REF. 1).
FT CONFLICT 867 867 A -> L (IN REF. 3).
SQ SEQUENCE 906 AA; 99851 MW; 72DDC7B8B57C7AFC CRC64;

Query Match 53.6%; Score 52; DB 1; Length 906;
Best Local Similarity 47.4%; Pred. No. 0.66;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFKIVSQEPA 19
:|||||: :|||:
Db 296 DPWALNGMLRYRILSQAPS 314

RESULT 8
CAD2_CRIGR STANDARD; PRT; 238 AA.
ID_CAD2_CRIGR
```

```

OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90360979; PubMed=2390969;
RA Llaw C.W., Cannon C., Power M.D., Kibonoka P.K., Rubin L.L.;
RT "Identification and cloning of two species of cadherins in bovine
RT endothelial cells.";
RL EMBO J. 9:2701-2708(1990).
CC -1- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.
CC They preferentially interact with themselves in a homophilic
CC manner in connecting cells; cadherins may thus contribute to the
CC sorting of heterogeneous cell types. N-cadherin may be involved in
CC neuronal recognition mechanism.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 5 cadherin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X53615; CAA37677.1; -.
DR PIR; S11693; IJBOCN.
DR HSPP; P15116; INCI.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR002333; Cadherin_C term.
DR Pfam; PF00028; Cadherin; 5.
DR Pfam; PF01049; Cadherin_C term; 1.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 5.
DR PROSITE; PS00232; CADHERIN 1; 3.
DR PROSITE; PS00268; CADHERIN 2; 5.
KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat.
FT NON TER 1
FT PROPEP <1 130
FT CHAIN 131 877
FT DOMAIN 131 695
FT TRANSMEM 696 717
FT DOMAIN 718 877
FT DOMAIN 131 238
FT DOMAIN 239 363
FT DOMAIN 364 468
FT DOMAIN 469 574
FT DOMAIN 575 685
FT DOMAIN 834 849
FT CARBOHYD 161 161
FT CARBOHYD 244 244
FT CARBOHYD 296 296
FT CARBOHYD 373 373
FT CARBOHYD 543 543
FT CARBOHYD 622 622
FT CARBOHYD 663 663
SQ SEQUENCE 877 AA; 96845 MW; 441B929ED871A249 CRC64;

Query Match 50.5%; Score 49; DB 1; Length 877;
Best Local Similarity 42.1%; Pred. No. 2.1;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 EPNHNSKIAFKIVSQEPA 19
Db 267 DPNALNGMLRYILISQAPS 285

RESULT 10
CAD2_MOUSE
ID CAD2_MOUSE STANDARD; PRT; 906 AA.
AC P15116; Q64260;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)

DT DE
GN NCBI_TaxID=9913;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89346748; PubMed=2762814;
RA Miyatani S., Shimamura K., Hatta M., Nagafuchi A., Nose A.,
RA Matsunaga M., Hatta K., Takeichi M.;
RT "Neural cadherin: role in selective cell-cell adhesion.";
RL Science 245:631-635(1989).
RN [2]
RP SEQUENCE FROM N.A.
RA Tamura K.;
RA Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP PARTIAL SEQUENCE FROM N.A.
RX STRAIN=C57BL/6;
RX MEDLINE=92409532; PubMed=1528849;
RA Miyatani S., Copeland N.G., Gilbert D.J., Jenkins N.A., Takeichi M.;
RT "Genomic structure and chromosomal mapping of the mouse N-cadherin
RT gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:8443-8447(1992).
RN [4]
RP DEVELOPMENTAL STAGE.
RX STRAIN=C57BL/6; TISSUE=Testis;
RX MEDLINE=97033837; PubMed=8879495;
RA Munro S.B., Blaschuk O.W.;
RT "A comprehensive survey of the cadherins expressed in the testes of
RT fetal, immature, and adult mice utilizing the polymerase chain
RT reaction.";
RL Biol. Reprod. 55:822-827(1996).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 160-267.
RX MEDLINE=95191680; PubMed=7885471;
RA Shapiro L., Fannon A.M., Kwong P.D., Thompson A., Lehmann M.S.,
RA Grubel G., Legrand J.-F., Als-Nielsen J., Colman D.R.,
RA Hendrickson W.A.;
RT "Structural basis of cell-cell adhesion by cadherins.";
RL Nature 374:327-337(1995).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (3.4 ANGSTROMS) OF 160-374.
RX MEDLINE=98318235; PubMed=9655503;
RA Tamura K., Shan W.S., Hendrickson W.A., Colman D.R., Shapiro L.;
RT "Structure-function analysis of cell adhesion by neural (N-)
RT cadherin.";
RL Neuron 20:1153-1163(1998).
CC -1- FUNCTION: Cadherins are calcium dependent cell adhesion proteins.
CC They preferentially interact with themselves in a homophilic
CC manner in connecting cells; cadherins may thus contribute to the
CC sorting of heterogeneous cell types. N-cadherin may be involved in
CC neuronal recognition mechanism.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- DEVELOPMENTAL STAGE: Expressed at all stages of testicular
CC development with highest levels found in testes of 21-day-old
CC mice.
CC -1- SIMILARITY: Contains 5 cadherin domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M31131; AAA37353.1; -.
DR EMBL; AB008811; BAA23549.1; -.
DR EMBL; S45011; AAB23356.1; -.
DR PIR; A32759; IJMSCN.

```



```
CC      membranes of renal tubular epithelial cells.
CC      -1- SIMILARITY: Contains 6 cadherin domains.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: U28945; AAC48472.1; -
CC      PIR: I46536; I46536.
CC      HSP: P09803; ISUH.
CC      InterPro: IPR002126; Cadherin.
CC      Pfam: PF00028; cadherin; 6.
CC      PRINTS: PR00205; CADHERIN.
CC      SMART: SM00112; CA; 6.
CC      PROSITE: PS00232; CADHERIN 1; 2.
CC      PROSITE: PS0268; CADHERIN 2; 6.
CC      Cell adhesion; Glycoprotein; transmembrane; Calcium-binding; Repeat;
CC      signal.
CC      FT SIGNAL. 1 18
CC      FT CHAIN. 19 829
CC      FT DOMAIN. 19 786
CC      FT EXTRACELLULAR (POTENTIAL).
CC      FT TRANSMEM. 787 807
CC      FT CYTOPLASMIC (POTENTIAL).
CC      FT DOMAIN. 808 829
CC      FT DOMAIN. 25 126
CC      FT DOMAIN. 131 235
CC      FT DOMAIN. 242 336
CC      FT DOMAIN. 341 449
CC      FT DOMAIN. 455 564
CC      FT DOMAIN. 569 665
CC      FT DOMAIN. 666 786
CC      FT CARBOHYD. 517 517
CC      FT CARBOHYD. 602 602
CC      FT CARBOHYD. 709 709
CC      FT CARBOHYD. 722 722
CC      FT CARBOHYD. 722 722
CC      SEQUENCE 829 AA; 88827 MW; D2DF10E6C47A43B9 CRC64;
CC
CC      Query Match 49.5%; Score 48; DB 1; Length 829;
CC      Best Local Similarity 52.6%; Pred. No. 2.9;
CC      Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
CC
CC      QY 1 BPHNSKIAPKIVSQEPA 19
CC      ||| : : : : :
CC      Db 155 EFGTANSDLRPHLSQTTPA 173
CC
CC      RESULT 14
CC      HIS1_PASMU STANDARD; PRT; 299 AA.
CC      AC P57919;
CC      DT 16-OCT-2001 (Rel. 40, Created)
CC      DT 28-FEB-2003 (Rel. 41, Last sequence update)
CC      DE ATP phosphoribosyltransferase (EC 2.4.2.17).
CC      GN HISG OR PML195.
CC      OS Pasteurella multocida.
CC      OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
CC      OC Pasteurellaceae; Pasteurella.
CC      OX NCBI_TaxID=747;
CC      RN [1]
CC      RP SEQUENCE FROM N.A.
CC      RC STRAIN=Fm70;
CC      RX MEDLINE=21145866; PubMed=11248100;
CC      RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
CC      "Complete genomic sequence of Pasteurella multocida Pm70.";
CC      Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC      -1- CATALYTIC ACTIVITY: 1-(5-phospho-D-ribose)-ATP + diphosphate =
CC      ATP + 5-phospho-alpha-D-ribose 1-diphosphate..
CC      -1- PATHWAY: Histidine biosynthesis; first step. Very important in the
CC      regulation of histidine metabolism.
```

```
CC      -1- SUBUNIT: Homohexamer (By similarity).
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC      -1- SIMILARITY: Belongs to the ATP phosphoribosyltransferase family.
CC      Long subfamily.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: AE006160; AAK03279.1; -
CC      HAMAP: MF_00079; -; 1.
CC      InterPro: IPR001348; ATP_phospho_trans.
CC      Pfam: PF01634; HisG; 1.
CC      ProDom: PD003516; ATP_phospho_trans; 1.
CC      TIGRFAMs: TIGR00070; HisG; 1.
CC      PROSITE: PS01316; ATP_P_PHORIBOSYLTR; 1.
CC      Histidine biosynthesis; Transferase; Glycosyltransferase;
CC      Complete proteome.
CC      SEQUENCE 299 AA; 33032 MW; 578AF47B4D3BDE96 CRC64;
CC
CC      Query Match 47.4%; Score 46; DB 1; Length 299;
CC      Best Local Similarity 64.3%; Pred. No. 2.2;
CC      Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
CC
CC      QY 4 HLNKSIAPKIVSQE 17
CC      ||| : : : : :
CC      Db 258 HDNSKVAMHVVSQE 271
CC
CC      RESULT 15
CC      CADH_MOUSE STANDARD; PRT; 827 AA.
CC      ID CADH_MOUSE
CC      AC Q9R100;
CC      DT 16-OCT-2001 (Rel. 40, Created)
CC      DT 16-OCT-2001 (Rel. 40, Last sequence update)
CC      DT 15-MAR-2004 (Rel. 43, Last annotation update)
CC      DE Cadherin-17 precursor (Liver-intestine-cadherin) (LI-cadherin) (BILL-
CC      cadherin) (P130).
CC      GN CDH17.
CC      OS Mus musculus (Mouse).
CC      OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC      OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC      OX NCBI_TaxID=10090;
CC      RN [1]
CC      RP SEQUENCE FROM N.A.
CC      RC STRAIN=Swiss Webster; TISSUE=Intestine;
CC      RX MEDLINE=21270055; PubMed=11376485;
CC      RA Angres B., Kim L., Jung R., Gessner R., Tauber R.;
CC      "Li-cadherin gene expression during mouse intestinal development.";
CC      Dev. Dyn. 221:182-193(2001).
CC      RN [2]
CC      RP SEQUENCE FROM N.A.; SEQUENCE OF 26-33; 52-58; 74-81; 117-123 AND
CC      RP 490-509. TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
CC      RC STRAIN=BAH/c; TISSUE=Fetal liver;
CC      RX MEDLINE=20469471; PubMed=10906147;
CC      RA Ohnishi K., Shimizu T., Karasuyama H., Melchers F.;
CC      "The identification of a nonclassical cadherin expressed during B cell
CC      development and its interaction with surrogate light chain.";
CC      J. Biol. Chem. 275:31134-31144(2000).
CC      RN [3]
CC      RP SEQUENCE FROM N.A.
CC      RC STRAIN=FVB/N; TISSUE=Colon;
CC      RX MEDLINE=22389257; PubMed=12477932;
CC      RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
CC      RA Strausberg R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
CC      RA Klausner R.D., Collins F.S., Wagner K.H., Schaefer C.F., Bhat N.K.,
CC      RA Altschul S.F., Zeeberg B., Buetow K.H., Max S.I., Wang J., Hsieh F.,
CC      RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
CC      RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
CC      Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2004, 05:56:44 ; Search time 66 Seconds
(without alignments)
90.831 Million cell updates/sec

Title: US-10-799-005A-1

Perfect score: 97

Sequence: 1 EPNHLSKIAFKIVSQEPA 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phase:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rvirus:*

16: sp_bacteriap:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97	100.0	993	6 Q7YRU7	Q7YRU7 canis famil
2	95	97.9	929	11 Q8CE48	Q8CE48 mus musculu
3	95	97.9	993	11 Q8CB02	Q8CB02 mus musculu
4	85	87.6	1041	11 Q7TMD7	Q7TMD7 mus musculu
5	83	85.6	1040	4 Q86S06	Q86S06 homo sapien
6	81	83.5	1054	6 Q8GK08	Q8GK08 canis famil
7	75	77.3	911	11 Q7TSF0	Q7TSF0 mus musculu
8	75	77.3	911	11 Q7TQ61	Q7TQ61 mus musculu
9	75	77.3	1060	11 Q7TSF1	Q7TSF1 mus musculu
10	75	77.3	1060	11 Q7TQ60	Q7TQ60 mus musculu
11	65	67.0	292	11 Q8VCE3	Q8VCE3 mus musculu
12	65	67.0	360	11 Q811I1	Q811I1 mus musculu
13	54	55.7	639	10 Q7YWF5	Q7YWF5 oryza sativ
14	54	55.7	864	10 Q7XVW3	Q7XVW3 oryza sativ
15	53	54.6	200	9 Q8SD26	Q8SD26 pseudomonas
16	52	53.6	533	2 Q54891	Q54891 streptococc

17	52	53.6	533	2 Q54899	Q54899 streptococc
18	52	53.6	533	16 Q8NZ78	Q8NZ78 streptococc
19	52	53.6	906	4 Q8N173	Q8N173 homo sapien
20	49	50.5	707	16 Q8S5J8	Q8S5J8 rhizobium 1
21	49	50.5	906	11 Q8BSI9	Q8BSI9 mus musculu
22	47.5	49.0	464	16 Q8QOR0	Q8QOR0 bradyrhizob
23	47	48.5	317	16 Q97LP1	Q97LP1 clostridium
24	46	47.4	334	2 Q9Z643	Q9Z643 enterobacte
25	45	46.4	459	16 Q82Y84	Q82Y84 nitrosomona
26	44	45.4	347	17 Q29274	Q29274 archaeoglob
27	44	45.4	441	5 Q97300	Q97300 plasmodium
28	44	45.4	484	11 Q8BG95	Q8BG95 mus musculu
29	44	45.4	512	3 Q13985	Q13985 schizosacch
30	44	45.4	1225	5 Q22656	Q22656 caenorhabdi
31	44	45.4	1269	5 Q22655	Q22655 caenorhabdi
32	44	45.4	1732	5 Q963J6	Q963J6 heliothis v
33	43.5	44.8	1928	5 Q8T9H1	Q8T9H1 drosophila
34	43.5	44.8	2280	5 Q9V8E6	Q9V8E6 drosophila
35	43.5	44.8	2302	5 Q9N693	Q9N693 drosophila
36	43.5	44.8	2310	5 Q9GRA9	Q9GRA9 drosophila
37	43	44.3	66	2 Q7WXC8	Q7WXC8 alicigenes
38	43	44.3	99	5 Q9UAD7	Q9UAD7 haemochus
39	43	44.3	99	5 Q9UAD8	Q9UAD8 haemochus
40	43	44.3	100	5 Q9UAD5	Q9UAD5 haemochus
41	43	44.3	100	5 Q9UAE3	Q9UAE3 haemochus
42	43	44.3	100	5 Q9UAD4	Q9UAD4 haemochus
43	43	44.3	100	5 Q9UAD6	Q9UAD6 haemochus
44	43	44.3	100	5 Q9UAD9	Q9UAD9 haemochus
45	43	44.3	272	16 Q889A5	Q889A5 pseudomonas

ALIGNMENTS

RESULT 1

Q7YRU7	Q7YRU7	PRELIMINARY;	PRT;	993 AA.
AC	Q7YRU7;			
DT	01-OCT-2003 (Tremblrel. 25, Created)			
DT	01-OCT-2003 (Tremblrel. 25, Last sequence update)			
DT	01-OCT-2003 (Tremblrel. 25, Last annotation update)			
DE	Desmoglein 3.			
GN	DSG3.			
OS	Canis familiaris (Dog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.			
OX	NCBI_TaxID=9615;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Oral mucosa;			
RA	Nishifuji K., Amagai M., Ota T., Park S.-J., Nishikawa T., Iwasaki T.;			
RL	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF394784; AAP80592.1; -			
SQ	SEQUENCE 993 AA; 107550 MW; B1E9C6DE9666D5D CRC64;			

Query Match 100.0%; Score 97; DB 6; Length 993;

Best Local Similarity 100.0%; Pred. No. 9e-08;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFKIVSQEPA 19

Db 185 EPNHLSKIAFKIVSQEPA 203

RESULT 2

Q8CE48

ID Q8CE48

AC Q8CE48

DT 01-MAR-2003 (Tremblrel. 23, Created)

DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)

DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)

DE Desmoglein 3.

GN DSG3.


```

ID Q7TSF1 PRELIMINARY; PRT; 1060 AA.
AC Q7TSF1;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Desmoglein 5.
DN DSG5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22672408;
RA Whittcock N.V.;
RT "Genetic analysis of the mouse desmoglein cluster reveals evidence for
RT six distinct genes: Characterization of Mouse DSG4, DSG5, and DSG6.";
RL J. Invest. Dermatol. 120:970-980(2003).
DR EMBL; AY192158; AAP31152.1; -.
DR EMBL; AY192158; AAP31152.1; -.
SQ SEQUENCE 1060 AA; 114453 MW; 58873A358C52B9C8 CRC64;

Query Match 77.3%; Score 75; DB 11; Length 1060;
Best Local Similarity 73.7%; Pred. No. 0.00058;
Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EPNHLSKIAPKIVSQEPA 19
DB 186 EPNHLSMIAFKIIRQEPS 204

RESULT 10
Q7TQ60 PRELIMINARY; PRT; 1060 AA.
ID Q7TQ60
AC Q7TQ60;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Desmoglein lbeta.
DN DSG1B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=FWB/N;
RC STRAIN=FWB/N;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC044848; AAH44848.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0007156; P:homophilic cell adhesion; IEA.
DR EXP. Dermatol. 12:11-19(2003).
DR EMBL; AY315940; AAP80571.1; -.
DR EMBL; AY315940; AAP80571.1; -.
SQ SEQUENCE 1060 AA; 114452 MW; F79BC83C9E474FA4 CRC64;

Query Match 77.3%; Score 75; DB 11; Length 1060;
Best Local Similarity 73.7%; Pred. No. 0.00058;
Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EPNHLSKIAPKIVSQEPA 19
DB 186 EPNHLSMIAFKIIRQEPS 204

RESULT 11
Q8VCE3 PRELIMINARY; PRT; 292 AA.
ID Q8VCE3
AC Q8VCE3;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein.
DN DSG2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast tumor;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC020144; AAH20144.1; -.
DR MGD; MGI:1196466; Dsg2.
DR GO; GO:0030057; C:desmosome; IEA.
DR InterPro; IPR002126; C:cadherin.
DR Pfam; PF00028; cadherin; 2.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 2.
DR PROSITE; PS00232; CADHERIN_1; 2.
DR PROSITE; PS0268; CADHERIN_2; 2.
KW Hypothetical protein; Calcium; Calcium-binding; Cell adhesion;
KW Glycoprotein.
SQ SEQUENCE 292 AA; 32850 MW; 7C2E26D76FA6FEAF CRC64;

Query Match 67.0%; Score 65; DB 11; Length 292;
Best Local Similarity 57.9%; Pred. No. 0.0079;
Matches 11; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 1 EPNHLSKIAPKIVSQEPA 19
DB 193 DPETLNKSVYRIVSQEPA 211

RESULT 12
Q811I1 PRELIMINARY; PRT; 360 AA.
ID Q811I1
AC Q811I1;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Similar to desmoglein 2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=FWB/N;
RC STRAIN=FWB/N;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC044848; AAH44848.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0007156; P:homophilic cell adhesion; IEA.
DR InterPro; IPR002126; Cadherin.
DR Pfam; PF00028; cadherin; 3.
DR PRINTS; PR00205; CADHERIN.
DR SMART; SM00112; CA; 2.
DR PROSITE; PS00232; CADHERIN_1; 2.
DR PROSITE; PS0268; CADHERIN_2; 3.
SQ SEQUENCE 360 AA; 40312 MW; 6A758B84705E8245 CRC64;

Query Match 67.0%; Score 65; DB 11; Length 360;
Best Local Similarity 57.9%; Pred. No. 0.0098;
Matches 11; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

QY 1 EPNHLSKIAPKIVSQEPA 19
DB 193 DPETLNKSVYRIVSQEPA 211

RESULT 13
Q7Y0F5 PRELIMINARY; PRT; 639 AA.
ID Q7Y0F5
AC Q7Y0F5;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)

```



```

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative polyprotein.
GN OSJNB0074M06.6.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN 1
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.B., Tsitrin T., Kim M.M., Bera J.J., Jin S.S.,
RA Fadrosh D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Uterback T.T., Feldblyum T.V.,
RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 3 BAC OSJNB0074M06 genomic sequence.";
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN 2
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Buell R.;
RA Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC133932; AAP50929.1; -.
KW Polyprotein.
SQ SEQUENCE 639 AA; 72114 MW; A2BCD632834F4288 CRC64;

Query Match 55.7%; Score 54; DB 10; Length 639;
Best Local Similarity 73.3%; Pred. No. 1.4;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EPNHNSKIAPKIVS 15
Db 218 EPNHNSLNAPKEIS 232

RESULT 14
Q7XVW3 PRELIMINARY; PRT; 864 AA.
AC Q7XVW3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE OSJNB0065J03.8 protein.
GN OSJNB0065J03.8.
OS Eukaryota; Viridiplantae;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN 1
RP SEQUENCE FROM N.A.
RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,
RA Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Weng Q.J., Zhang L.,
RA Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T.,
RA Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.,
RA Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,
RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,
RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,
RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,
RA Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL731615; CAD40412.1; -.
SQ SEQUENCE 864 AA; 97804 MW; 68AA1554147D79A6 CRC64;

Query Match 55.7%; Score 54; DB 10; Length 864;
Best Local Similarity 73.3%; Pred. No. 1.9;
Matches 11; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EPNHNSKIAPKIVS 15
Db 162 EPNHNSLNAPKEIS 176

```

```

RESULT 15
Q8SD26 PRELIMINARY; PRT; 200 AA.
AC Q8SD26;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE PHK2136.
OS Pseudomonas phage phiK2.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
OX NCBI_TaxID=169683;
RN 1
RP SEQUENCE FROM N.A.
RX MEDLINE=21914557; PubMed=11916376;
RA Mesyanzhinov V.V., Robben J., Grymonprez B., Kostyuchenko V.A.,
RA Bourkaltseva M.V., Sykilinda N.N., Krylov V.V., Volckaert G.;
RA Bourkaltseva M.V., Sykilinda N.N., Krylov V.N., Volckaert G.;
RT "The genome of bacteriophage phiK2 of Pseudomonas aeruginosa.";
RL J. Mol. Biol. 317:1-19 (2002).
RN 2
RP SEQUENCE FROM N.A.
RA Mesyanzhinov V.V., Robben J., Grymonprez B., Kostyuchenko V.A.,
RA Bourkaltseva M.V., Sykilinda N.N., Krylov V.V., Volckaert G.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF399011; AAL83037.1; -.
SQ SEQUENCE 200 AA; 22757 MW; B8C22293FD79494E CRC64;

Query Match 54.6%; Score 53; DB 9; Length 200;
Best Local Similarity 56.2%; Pred. No. 0.61;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 3 NHLNSKIAPKIVSQEP 18
Db 96 DHLNSKVAYKITBYP 111

```

Search completed: October 12, 2004, 06:00:14
Job time : 73 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2004, 05:56:44 ; Search time 71 Seconds
(without alignments)
75.611 Million cell updates/sec

Title: US-10-799-005A-1
Perfect score: 97
Sequence: 1 EPNHLSKIAFKIVSQEPA 19

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97	100.0	614	2 AAW07908	Aaw07908 Pemphigus
2	97	100.0	999	2 AAR30742	Aar30742 Human pem
3	97	100.0	999	5 AAU78054	Aau78054 Human des
4	97	100.0	999	6 ABR48239	Abu48239 Human bla
5	97	100.0	999	6 ABUS6419	Abu56419 Lung canc
6	97	100.0	999	6 ABUS6662	Abu56662 Lung canc
7	97	100.0	999	6 ADA83744	Ada83744 Human DSG
8	97	100.0	1014	4 ABG12435	Abg12435 Novel hum
9	83	85.6	274	4 ABG12434	Abg12434 Novel hum
10	83	85.6	1191	4 ABG12436	Abg12436 Novel hum
11	81	83.5	112	3 AAY64603	Aay64603 Nonclassi
12	81	83.5	778	2 AAW15489	Aaw15489 Pemphigus
13	81	83.5	1049	5 AAU78052	Aau78052 Human des
14	71	73.2	15	2 AAW04843	Aaw04843 Self epit
15	71	73.2	15	2 AAW64815	Aaw64815 Desmoglei
16	71	73.2	15	2 AAW78814	Aaw78814 Desmoglei
17	71	73.2	15	3 AAB33625	Aab33625 MHC class
18	71	73.2	15	4 AAG93721	Aag93721 Human des
19	71	73.2	15	5 AAOL17033	Aao17033 Desmoglei
20	71	73.2	15	6 ABU96577	Abu96577 MHC class
21	71	73.2	113	3 AAY64604	Aay64604 Nonclassi
22	71	73.2	263	2 AAW13010	Aaw13010 Segment o
23	71	73.2	560	2 AAW13009	Aaw13009 Segment o
24	71	73.2	940	4 AAM39436	Aam39436 Human pol
25	71	73.2	1117	5 AAU78053	Aau78053 Human des

26	71	73.2	1118	5 AAE20266	Aae20266 Human lun
27	71	73.2	1120	4 ABG10284	Abg10284 Novel hum
28	71	73.2	1121	4 AAM41222	Aam41222 Human pol
29	52	53.6	747	7 ADD14035	Add14035 Human src
30	52	53.6	848	5 ABG79689	Abg79689 Tumour in
31	52	53.6	906	3 AAY70741	Aay70741 Human N-c
32	52	53.6	906	4 ABG14316	Abg14316 Novel hum
33	52	53.6	906	5 ABB81474	Abb81474 Human N-c
34	52	53.6	906	6 ABR58643	Abt58643 Human can
35	52	53.6	906	6 ABR47406	Abt47406 Breast ca
36	52	53.6	906	7 ADE55478	Ade55478 Human pro
37	52	53.6	906	7 ADE55482	Ade55482 Human pro
38	52	53.6	906	7 ADE55486	Ade55486 Human pro
39	52	53.6	906	7 ADE55490	Ade55490 Human pro
40	49	50.5	906	5 ABE57233	Abb57233 Mouse isc
41	49	50.5	906	7 ADE55488	Ade55488 Rat prote
42	49	50.5	906	7 ADE55476	Ade55476 Rat prote
43	49	50.5	906	7 ADE55484	Ade55484 Rat prote
44	49	50.5	906	7 ADE55480	Ade55480 Rat prote
45	48	49.5	807	3 AAY99405	Aay99405 Human pro

ALIGNMENTS

RESULT 1
AAW07908
ID AAW07908 standard; protein; 614 AA.

XX AC AAW07908;

XX DT 29-JAN-1997 (first entry)

XX DE Pemphigus vulgaris antigen protein extracellular region.

XX KW Autoantibody; immunoglobulin G; IgG1; fusion protein; diagnosis;
treatment; pemphigus vulgaris; PV; Bulla; blister; skin disease;
dermatology.

XX OS Homo sapiens.

XX PN JP08188540-A.

XX PD 23-JUL-1996.

XX PF 30-JUN-1995; 95JP-00165632.

XX PR 30-JUN-1994; 94JP-00173291.

XX PA (NISH/) NISHUKAWA T.

XX DR WPI; 1996-388562/39.

XX PT Fused protein recognised by Pemphigus vulgaris auto-antibody - useful to
treat and diagnose P. vulgaris related diseases.

XX PS Claim 1; Page 7-9; 9pp; Japanese.

XX CC AAW07908 represents the human pemphigus vulgaris (PV) antigen
extracellular region. The PV antigen is produced in patients with pemphig
vulgaris resulting in autoimmune disease. PV is a rare relapsing disease
causing suprabasal, intra-epidermal bullae (vesicles) of the skin and
mucous membranes, which is fatal if untreated. The PV antigen was fused
to a human IgG1 hinge region and the resulting fusion protein is useful
to treat or diagnose pemphigus vulgaris

XX SQ Sequence 614 AA;

Query Match 100.0%; Score 97; DB 2; Length 614;
Best Local Similarity 100.0%; Pred. No. 3.1e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFKIVSQEPA 19

Human; desmoglein; stratified squamous epithelia; noxious substance; extracellular domain; junctional protein; luminal damage; heartburn; laryngitis; gastric influx; gastroesophageal reflux disease; GERD; sore throat; pemphigus vulgaris antigen.

XX 03-JUL-2002; 2002WO-US021338.
 XX 03-JUL-2001; 2001US-0302814P.
 PR 03-AUG-2001; 2001US-0310099P.
 PR 08-NOV-2001; 2001US-0343705P.
 PR 13-NOV-2001; 2001US-0350666P.
 PR 12-APR-2002; 2002US-0372246P.
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 XX Mack DH, Aziz N;
 XX WPI; 2003-201532/19.
 DR N-PSDB; ACC51055.
 XX Detecting a bladder cancer-associated transcript in a cell from a
 PT patient, comprises contacting a biological sample from the patient with a
 PT bladder cancer-associated polynucleotide or antibody.
 XX Claim 10; Page 299; 307pp; English.
 PS The present invention describes a method for detecting a bladder cancer-
 XX associated transcript in a cell from a patient. The method comprises
 CC contacting a biological sample from the patient with a polynucleotide
 CC that selectively hybridizes to a sequence that is 80 % identical to a
 CC table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059
 CC encode the human bladder cancer-associated proteins given in ABR48146 to
 CC ABR48242). Bladder cancer-associated proteins from the present invention
 CC have cytostatic activities, and can be used in antisense gene therapy and
 CC in vaccine production. The method can be used for detecting a bladder
 CC cancer-associated transcript in a cell from a patient. The method is
 CC useful in diagnosing or treating bladder cancer and in screening for
 CC compounds that modulate bladder cancer, such as hormones or antibodies.
 CC The nucleic acid molecules from the present invention may be used in
 CC various screening and diagnostic methods, and for gene therapy, vaccine
 CC and/or antisense/inhibition applications
 XX Sequence 999 AA;
 SQ Query Match 100.0%; Score 97; DB 6; Length 999;
 Best Local Similarity 100.0%; Pred. No. 5.5e-08;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EPNHLSKIAFKIVSQEPA 19
 Db 186 EPNHLSKIAFKIVSQEPA 204
 RESULT 5
 ABUS6419
 ID ABUS6419 standard; protein; 999 AA.
 XX AC ABUS6419;
 XX DT 02-APR-2003 (first entry)
 DE Lung cancer-associated polypeptide #12.
 KW Lung cancer-associated polypeptide; cytostatic; emphysema;
 KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
 XX Unidentified.
 OS WO200286443-A2.
 XX 31-OCT-2002.
 XX 18-APR-2002; 2002WO-US012476.

PR 18-APR-2001; 2001US-0284770P.
 PR 10-MAY-2001; 2001US-0290492P.
 PR 09-NOV-2001; 2001US-0339245P.
 PR 13-NOV-2001; 2001US-0350666P.
 PR 29-NOV-2001; 2001US-0334370P.
 PR 12-APR-2002; 2002US-0372246P.
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 XX Aziz N, Murray R;
 XX WPI; 2003-093161/08.
 DR N-PSDB; ABX76135.
 XX Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer.
 XX Claim 27; Page 197; 453pp; English.
 XX The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridizes
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences
 CC ABUS6408-ABUS6745 represent lung cancer-associated polypeptides of the
 CC invention
 XX Sequence 999 AA;
 SQ Query Match 100.0%; Score 97; DB 6; Length 999;
 Best Local Similarity 100.0%; Pred. No. 5.5e-08;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 EPNHLSKIAFKIVSQEPA 19
 Db 186 EPNHLSKIAFKIVSQEPA 204
 RESULT 6
 ABUS6662
 ID ABUS6662 standard; protein; 999 AA.
 XX AC ABUS6662;
 XX DT 02-APR-2003 (first entry)
 DE Lung cancer-associated polypeptide #255.
 KW Lung cancer-associated polypeptide; cytostatic; emphysema;
 KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
 XX Unidentified.
 OS WO200286443-A2.
 XX 31-OCT-2002.

XX PF 18-APR-2002; 2002WO-US012476.
XX PR 18-APR-2001; 2001US-0284770P.
XX PR 10-MAY-2001; 2001US-0290492P.
XX PR 09-NOV-2001; 2001US-0339245P.
XX PR 13-NOV-2001; 2001US-0350666P.
XX PR 29-NOV-2001; 2001US-0334370P.
XX PR 12-APR-2002; 2002US-0372246P.
XX PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX PI Aziz N, Murray R;
XX DR WPI; 2003-093161/08.
XX DR N-PSDB; ABX76391.
XX PT Detecting a lung cancer-associated transcript in a cell from a patient
XX PT for treating lung cancer, by contacting a biological sample from the
XX PT patient with a polynucleotide that exhibits increased or decreased
XX PT expression in lung cancer.
XX PS Claim 27; Page 385; 453pp; English.
XX CC The invention relates to a method for detecting a lung cancer-associated
XX CC transcript in a cell from a patient, comprising contacting a biological
XX CC sample from the patient with a polynucleotide that selectively hybridises
XX CC to a sequence that is at least 80 % identical to a gene that exhibits
XX CC increased or decreased expression in lung cancer samples. Lung cancer-
XX CC associated polynucleotides and polypeptides are used for identifying a
XX CC compound that modulates a lung cancer-associated polypeptide, for
XX CC inhibiting proliferation of a lung cancer-associated cell to treat lung
XX CC cancer in a patient and for treating a mammal having lung cancer by
XX CC administering a modulatory compound identified. The methods are useful
XX CC for treating lung cancer, such as small cell lung cancer, non-small cell
XX CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
XX CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
XX CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
XX CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
XX CC for diagnostic purposes and as targets for screening for therapeutic
XX CC compounds that modulate lung cancer, such as antibodies. Sequences
XX CC ABU56408-ABU56745 represent lung cancer-associated polypeptides of the
XX CC invention
XX SQ Sequence 999 AA;
Query Match 100.0%; Score 97; DB 6; Length 999;
Best Local Similarity 100.0%; Pred. No. 5.5e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EPNHLNSKIAPKIVSQEPA 19
| | | | | | | | | | | | | | | | | | | | | |
Db 186 EPNHLNSKIAPKIVSQEPA 204
RESULT 7
ADA83744
ID ADA83744 standard; protein; 999 AA.
XX AC ADA83744;
XX DT 20-NOV-2003 (first entry)
XX DE Human DSG3 protein.
XX KW human; marker; expressed sequence tag; EST; arabadopsis; tumour;
XX KW stress-induced phenotype; hyperosmotic stress; colon cancer; immunogen;
XX KW vaccine.
XX OS Homo sapiens.
XX PN WO2002103028-A2.
XX PF

PD XX 27-DEC-2002.
XX PF 30-MAY-2002; 2002WO-IB004189.
XX PR 30-MAY-2001; 2001US-0293999P.
XX PR 22-OCT-2001; 2001US-0330457P.
XX PR 19-FEB-2002; 2002US-0357144P.
XX PA (BIOM-) BIOMEDICAL CENT.
XX PI Baranova AV, Yankovsky NK, Kozlov AP, Lobashev AV, Krukovskaya LL;
XX DR WPI; 2003-175241/17.
XX DR N-PSDB; ADA83743.
XX PT Determining if a nucleic acid is a marker for a phenotype/cell type of
XX PT interest, by global comparison of expressed sequence tags known to be
XX PT expressed in the phenotype/cell type with all ESTs expressed in normal
XX PT tissue.
XX PS Claim 29; Page 96-101; 516pp; English.
XX CC The invention relates to a novel method for determining if a nucleic acid
XX CC is a marker for a predetermined phenotype/cell type of interest from a
XX CC biological species. The method comprises performing a global comparison
XX CC of a group of expressed sequence tags (ESTs) known to be expressed in the
XX CC phenotype/cell type of interest with all ESTs expressed in normal tissue
XX CC in order to identify ESTs that are preferentially expressed in the
XX CC phenotype/cell of interest. A method of the invention is useful for
XX CC determining whether a nucleic acid is a marker for a predetermined
XX CC Arabidopsis or human. The cell type of interest is an abnormal cell such
XX CC as a tumour cell, and the predetermined phenotype is a stress-induced
XX CC phenotype such as hyperosmotic stress or high salt conditions. A method
XX CC of the invention is also useful for determining the progression of colon
XX CC cancer in a human, for detecting a tumour cell, and for regulating or
XX CC preventing the growth of a tumour cell. An antibody of the invention is
XX CC useful for detecting the absence or presence of peptides encoded by
XX CC tumour-associated markers. A polypeptide of the invention is useful as an
XX CC immunogen for vaccinating an animal. The present sequence represents a
XX CC tumour-associated antigen of the invention.
XX SQ Sequence 999 AA;
Query Match 100.0%; Score 97; DB 6; Length 999;
Best Local Similarity 100.0%; Pred. No. 5.5e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 EPNHLNSKIAPKIVSQEPA 19
| | | | | | | | | | | | | | | | | | | | | |
Db 186 EPNHLNSKIAPKIVSQEPA 204
RESULT 8
ABG12435
ID ABG12435 standard; protein; 1014 AA.
XX AC ABG12435;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #12426.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.

```

XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
XX DR N-PSDB; AAS76622.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX PS Claim 20; SEQ ID NO 42794; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX CC sequences. (I) is useful as hybridisation probes, polymerase chain
XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX CC and in recombinant production of (II). The polynucleotides are also used
XX CC in diagnostics as expressed sequence tags for identifying expressed
XX CC genes. (I) is useful in gene therapy techniques to restore normal
XX CC activity of (II) or to treat disease states involving (II). (II) is
XX CC useful for generating antibodies against it, detecting or quantitating a
XX CC polypeptide in tissue, as molecular weight markers and as a food
XX CC supplement. (II) and its binding partners are useful in medical imaging
XX CC of sites expressing (II). (I) and (II) are useful for treating disorders
XX CC involving aberrant protein expression or biological activity. The
XX CC polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
XX CC amino acid sequences of the invention. Note: The sequence data for this
XX CC patent did not appear in the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 1014 AA;

Query Match 100.0%; Score 97; DB 4; Length 1014;
Best Local Similarity 100.0%; Pred. No. 5.6e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EPNHLSKIAFKIVSQEPA 19
Db 201 EPNHLSKIAFKIVSQEPA 219

RESULT 9
ABG12434
ID ABG12434 standard; protein; 274 AA.
XX AC ABG12434;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #12425.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX FN 11-OCT-2001.
XX PD 30-MAR-2001; 2001WO-US008631.
XX PF 31-MAR-2000; 2000US-00540217.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.

XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
XX DR N-PSDB; AAS76621.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX PS Claim 20; SEQ ID NO 42793; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX CC sequences. (I) is useful as hybridisation probes, polymerase chain
XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX CC and in recombinant production of (II). The polynucleotides are also used
XX CC in diagnostics as expressed sequence tags for identifying expressed
XX CC genes. (I) is useful in gene therapy techniques to restore normal
XX CC activity of (II) or to treat disease states involving (II). (II) is
XX CC useful for generating antibodies against it, detecting or quantitating a
XX CC polypeptide in tissue, as molecular weight markers and as a food
XX CC supplement. (II) and its binding partners are useful in medical imaging
XX CC of sites expressing (II). (I) and (II) are useful for treating disorders
XX CC involving aberrant protein expression or biological activity. The
XX CC polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
XX CC amino acid sequences of the invention. Note: The sequence data for this
XX CC patent did not appear in the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 274 AA;

Query Match 85.6%; Score 83; DB 4; Length 274;
Best Local Similarity 84.2%; Pred. No. 3.5e-06;
Matches 16; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EPNHLSKIAFKIVSQEPA 19
Db 166 EENHLSKIAYKIVSQEPS 184

RESULT 10
ABG12436
ID ABG12436 standard; protein; 1191 AA.
XX AC ABG12436;
XX DT 18-FEB-2002 (first entry)
XX DE Novel human diagnostic protein #12427.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX FN 11-OCT-2001.
XX PD 30-MAR-2001; 2001WO-US008631.
XX PF 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PR 23-AUG-2000; 2000US-00649167.

```

```

PA (HYSE-) HYSEQ INC.
XX
XX Dmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX N-PSDB; AAS76623.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX
XX Claim 20; SEQ ID NO 42795; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
XX amino acid sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 1191 AA;
XX
XX Query Match 85.6%; Score 83; DB 4; Length 1191;
XX Best Local Similarity 84.2%; Pred. No. 2.1e-05; Indels 0; Gaps 0;
XX Matches 16; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 EPNHLSKIAFKIVSQEPA 19
XX |||||:|||||:|||||:
XX 1074 EBNHLSKIAYKIVSQEPS 1092
XX
XX RESULT 11
XX AAY64603
XX ID AAY64603 standard; peptide; 112 AA.
XX
XX AC AAY64603;
XX
XX DT 02-MAR-2000 (first entry)
XX
XX DE Nonclassical cadherin extracellular domain SEQ ID NO:31.
XX
XX KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;
XX inhibition; cadherin extracellular domain; cell adhesion recognition;
XX OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;
XX cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
XX cadherin related neuronal receptor; LI-cadherin; protocadherin;
XX desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
XX rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
XX neurological disease.
XX
XX OS Mammalia.
XX
XX PN WO9957149-A2.
XX
XX XX 11-NOV-1999.
XX
XX PD
XX XX 05-MAY-1999; 99WO-CA000363.
XX PF

```

```

XX 05-MAY-1998; 98US-00073040.
XX 06-NOV-1998; 98US-00187859.
XX 20-JAN-1999; 99US-00234395.
XX 08-MAR-1999; 99US-00264516.
XX
XX (ADHE-) ADHEREX TECHNOLOGIES INC.
XX
XX Blaschuk OW, Gour BJ, Byers S;
XX
XX WPI; 2000-038791/03.
XX
XX New cadherin modulating agents, used for modulating nonclassical cadherin
XX mediated functions for treating e.g. cancers, obesity, rheumatoid
XX arthritis, multiple sclerosis, diabetes or a neurological disease.
XX
XX Disclosure; Fig 2; 252pp; English.
XX
XX The present invention describes cadherin modulating agents (MA)
XX comprising peptides which comprise a nonclassical cadherin cell adhesion
XX recognition (CAR) sequence. The MAs can be used for modulating
XX nonclassical cadherin-mediated functions. They can be used for e.g.
XX inhibiting adhesion of nonclassical-cadherin expressing cells in a
XX mammal, enhancing delivery of a drug through the skin of a mammal,
XX enhancing delivery of a drug to a tumour in a mammal, treating cancer in
XX a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
XX angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
XX expressing cell, preventing or treating obesity in a mammal, stimulating
XX blood vessel regression in a mammal, enhancing drug delivery to the
XX central nervous system, treating a demyelinating neurological disease,
XX increasing vasopermeability in a mammal, enhancing adhesion of
XX nonclassical cadherin-expressing cells, inhibiting synaptic stability in
XX a mammal, or preventing pregnancy in a mammal. They can also be used for
XX e.g. enhancing or directing neurite outgrowth, facilitating wound healing
XX or reducing scar tissue, or enhancing adhesion of foreign tissue in a
XX mammal. They can also be used for treating e.g. psoriasis, arthritis, age
XX related macular degeneration, multiple sclerosis and diabetes. The
XX products can also be used for detection and diagnosis and in bioreactors.
XX AAY60592 to AAY64572 represent specifically claimed peptides, and
XX AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent sequences used in
XX the exemplification of the present invention
XX
XX SQ Sequence 112 AA;
XX
XX Query Match 83.5%; Score 81; DB 3; Length 112;
XX Best Local Similarity 78.9%; Pred. No. 2.7e-06;
XX Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 EPNHLSKIAFKIVSQEPA 19
XX |||||:|||||:|||||:
XX 29 EBNHLSKIAFKIIRQEPS 47
XX
XX Db
XX
XX RESULT 12
XX AAW15489
XX ID AAW15489 standard; protein; 779 AA.
XX
XX AC AAW15489;
XX
XX XX 17-OCT-2003 (revised)
XX DT 17-JUN-1997 (first entry)
XX
XX DE Pemphigus foliaceus antigen-IgG constant region fusion protein.
XX
XX KW Pemphigus foliaceus; autoantibody; constant region; IgG;
XX extracellular region; antigen; hinge portion; skin;
XX dermatitis herpetiformis; fusion protein; detection; ss.
XX
XX OS Homo sapiens.
XX
XX OS Chimeric.
XX
XX XX Key
XX FT Domain 1..545

```


/note= "Pemphigus foliaceus antigen protein"

FT XX JP09077800-A.
 PN XX
 XX XX
 PD XX
 XX XX
 XX XX
 PF 12-SEP-1995; 95JP-00260899.
 XX XX
 PR 12-SEP-1995; 95JP-00260899.
 XX XX
 XX (NISH/) NISHIKAWA T.
 PA XX
 XX WPI; 1997-241758/22.
 DR P-PSDB; AAT66428.
 XX XX
 XX Pemphigus foliaceus antigen-IgG constant region fusion protein - linked
 PT through the hinge region used to treat pemphigus foliaceus.
 XX XX
 PS Claim 1; Page 10-12; 17pp; Japanese.
 XX XX
 CC This sequence represents a fused protein recognised by pemphigus
 CC foliaceus patient autoantibody which comprises the constant region of IgG
 CC linked to the extracellular region of pemphigus foliaceus antigen protein
 CC through the hinge portion. Pemphigus foliaceus is a chronic, generalised,
 CC vesicular and scaling skin eruption similar to dermatitis herpetiformis.
 CC The pemphigus foliaceus antigen fusion protein is useful to treat
 CC pemphigus foliaceus. The antigen is especially administered through an
 CC adsorbent upon which the fusion protein is immobilised via a carrier. The
 CC fusion protein is also useful for detecting pemphigus foliaceus
 CC antibodies which is useful in immunodiagnosis. The fusion protein has
 CC little or no side effects. (Updated on 17-OCT-2003 to standardise OS
 CC field)
 XX XX
 SQ Sequence 778 AA;

Query Match 83.5%; Score 81; DB 2; Length 778;
 Best Local Similarity 78.9%; Pred. No. 2.9e-05;
 Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 EPNHLSKIAPKIVSQEPA 19
 |||:|||||||: |||:
 DB 186 EPNHLSKIAPKIVSQEPS 204

RESULT 13
 AAU78052
 ID AAU78052 standard; protein; 1049 AA.
 XX XX
 AC AAU78052;

DT 05-JUN-2002 (first entry)
 XX XX
 XX Human desmoglein 1 protein sequence.
 DE XX
 XX Human; desmoglein; stratified squamous epithelia; noxious substance;
 KW extracellular domain; junctional protein; luminal damage; heartburn;
 KW laryngitis; gastric influx; gastroesophageal reflux disease; GERD;
 KW sore throat.
 XX XX
 OS Homo sapiens.

XX Key Location/Qualifiers
 FH 50. .548
 FT Domain /label= Extracellular_domain
 FT Region 52. .157
 FT /note= "Specifically claimed in claim 8"

XX WO200210767-A2.
 XX PN
 XX 07-FEB-2002.
 XX PD
 XX 26-JUL-2001; 2001WO-US023717.
 XX PF
 XX XX

PR 28-JUL-2000; 2000US-00626196.
 XX XX
 XX (TULA) TULANE EDUCATIONAL FUND.
 XX PI
 XX Tobey NA, Orlando RC;
 XX DR
 XX WPI; 2002-257387/30.

XX Identifying compounds that protect stratified squamous epithelial tissue
 PT against injury by noxious substances, by determining interaction between
 PT a test compound and an E-cadherin extracellular domain.
 XX XX

PS Claim 7; Page 48-51; 62pp; English.

XX The present invention relates to a new method of assaying for a compound
 CC which may protect stratified squamous epithelia from damage by a noxious
 CC substance. The method of the invention involves determining the level or
 CC presence of an interaction between the test compound and a polypeptide
 CC sequence comprising a portion of the extracellular domain of the
 CC junctional protein E-cadherin or a related polypeptide sequence. The
 CC method is useful for identifying compounds, which may protect stratified
 CC squamous epithelial tissue against injury by noxious substances. In
 CC particular, the method is useful for identifying compounds for protective
 CC effects against luminal damaging compounds or compounds that may be
 CC useful in treating conditions associated with gastric reflux e.g.
 CC gastroesophageal reflux disease (GERD), or in treating sore throat,
 CC heartburn or laryngitis. The present amino acid sequence represents the
 CC human desmoglein 1 protein of the invention
 XX XX

SQ Sequence 1049 AA;

Query Match 83.5%; Score 81; DB 5; Length 1049;
 Best Local Similarity 78.9%; Pred. No. 4.1e-05;
 Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 EPNHLSKIAPKIVSQEPA 19
 |||:|||||||: |||:
 DB 186 EPNHLSKIAPKIVSQEPS 204

RESULT 14
 AAU04843
 ID AAU04843 standard; peptide; 15 AA.
 XX XX
 AC AAU04843;

DT 18-FEB-1997 (first entry)
 XX XX
 XX Self epitope of desmoglein 3, implicated in autoimmune disease.

DE Tolerisation; self-epitope; antigen; autoimmune disease; autoantigen;
 KW HLA; human leukocyte antigen; T-cell; thymocyte; pemphigus vulgaris;
 KW desmoglein; multiple sclerosis; herpes simplex virus; adenovirus;
 KW phosphomannomutase; human papillomavirus; Epstein-Barr virus;
 KW DNA polymerase; influenza; haemagglutinin; reovirus; sigma protein.

XX Homo sapiens.

XX WO9627387-A1.

XX 12-SEP-1996.

XX 07-MAR-1996; 96WO-US003182.

XX 07-MAR-1995; 95US-00400796.

XX (HARD) HARVARD COLLEGE.

XX Strominger JL, Wucherpfennig KW;

XX WPI; 1996-425218/42.

XX Pemphigus vulgaris auto-antigens and multiple sclerosis non-self antigens

PT - useful in disease treatment, and method for identification of other

PT self and non-self antigens implicated in auto-immune disease.

XX

XX Claim 1; Page 39; 58pp; English.

XX Pharmaceutical preparations for tolerisation to antigens comprise either

CC an isolated human non-collagen or non-mysin basic protein (MBP)

CC polypeptide which is capable of tolerising an individual to an

CC autoantigen; or an isolated human pathogen polypeptide capable of

CC tolerising an individual to that polypeptide. In both cases, the

CC polypeptide (whether self or non-self) includes an amino acid sequence

CC corresponding to a sequence motif for a MHC class II protein, such as HLA

CC -DR, which is associated with a human autoimmune disease and which binds

CC to the polypeptide to activate autoreactive T-cells in individuals with

CC the autoimmune disease. This peptide is derived from the human desmoglein

CC 3 protein (amino acids 190-204) and is implicated as a self epitope in

CC pemphigus vulgaris. Peptides derived from the human desmoglein protein

CC are described in AAW04841-47

XX

SQ Sequence 15 AA;

Query Match 73.2%; Score 71; DB 2; Length 15;

Best Local Similarity 100.0%; Pred. No. 1.4e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 LNSKIAPFKIVSQEPA 19

|||||

Db 1 LNSKIAPFKIVSQEPA 15

|||||

Search completed: October 12, 2004, 05:58:09

Job time : 77 secs

CC sequence, where the expression product of the coding sequence is a

CC polypeptide having a length and a sequence which permits it to bind to an

CC MHC Class I or II molecule. The expression product is thus an effective

CC stimulator of an immune response in mammals. The present sequence, an

CC antigenic portion of desmoglein 3, is an example of an MHC class II

CC peptide which can be expressed by the nucleic acid. It is associated with

XX pemphigus vulgaris

SQ Sequence 15 AA;

Query Match 73.2%; Score 71; DB 2; Length 15;

Best Local Similarity 100.0%; Pred. No. 1.4e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 LNSKIAPFKIVSQEPA 19

|||||

Db 1 LNSKIAPFKIVSQEPA 15

|||||

Search completed: October 12, 2004, 05:58:09

Job time : 77 secs

RESULT 15

AAW64815

ID AAW64815 standard; peptide; 15 AA.

XX

AC AAW64815;

DT 29-SEP-1998 (first entry)

XX

DE Desmoglein-3 190-204.

XX

KW Desmoglein; DG; gene therapy; pemphigus vulgaris; microparticle;

KW autoantigen; autoimmune disease; MHC.

OS Homo sapiens.

XX

PN US5783567-A.

XX

PD 21-JUL-1998.

XX

PF 22-JAN-1997; 97US-00787547.

XX

PR 22-JAN-1997; 97US-00787547.

XX

PA (PANG-) PANGAEA PHARM INC.

XX

PI Langer RS, Hedley ML, Curley JM;

XX

DR WPI; 1998-427077/36.

XX

PT Microparticle encapsulated nucleic acids - for recombinant expression of

PT proteins e.g. in gene therapy.

XX

PS Disclosure; Col 4; 42pp; English.

XX

CC The patent describes a new preparation of microparticles each comprising

CC a polymeric matrix and a nucleic acid. The polymeric matrix consists of

CC one or more synthetic polymers having a solubility in water of less than

CC 1 mg/l (e.g. poly-lactic-co-glycolic acid); and at least 90% of the

CC microparticles have a diameter of less than 100 microns. The

CC microparticles are useful for the delivery of nucleic acids to phagocytic

CC cells. In one embodiment the microparticles are less than 20 microns in

CC diameter and the nucleic acid (preferably in closed circular form)

CC includes an expression control sequence operatively linked to a coding

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2004, 05:58:38 ; Search time 71 Seconds
(without alignments)
86.115 Million cell updates/sec

Title: US-10-799-005A-1

Perfect score: 97

Sequence: 1 EPNHLNGKIAFKIVSQEPA 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	97	100.0	999	12	US-10-187-975-42
2	97	100.0	999	14	Sequence 42, Appl
3	97	100.0	999	15	Sequence 26, Appl
4	97	100.0	999	16	Sequence 322, App
5	81	83.5	112	14	US-10-188-832-199
6	81	83.5	112	15	US-10-006-869-31
7	71	73.2	115	9	US-10-395-032-31
8	71	73.2	15	2	US-09-909-460-24
9	71	73.2	113	14	US-09-872-836-24
10	71	73.2	113	15	US-10-006-869-32
11	71	73.2	1118	10	US-10-395-032-32
12	52	53.6	413	9	US-09-909-5678-42
13	52	53.6	848	12	Sequence 43, Appl
14	52	53.6	906	9	US-09-805-020-56
15	52	53.6	906	9	Sequence 56, Appl
					Sequence 46, Appl
					Sequence 46, Appl

16	52	53.6	906	14	US-10-177-293-43	Sequence 43, Appl
17	52	53.6	906	14	US-10-174-677-19	Sequence 19, Appl
18	52	53.6	906	16	US-10-778-146-6	Sequence 6, Appl
19	49	50.5	556	10	US-09-880-573-98	Sequence 98, Appl
20	49	50.5	877	9	US-09-746-491-47	Sequence 47, Appl
21	49	50.5	906	9	US-09-746-491-48	Sequence 48, Appl
22	48	49.5	807	10	US-09-946-374-229	Sequence 229, App
23	48	49.5	807	12	US-10-219-535-188	Sequence 188, App
24	48	49.5	807	12	US-10-232-230-188	Sequence 188, App
25	48	49.5	807	12	US-10-063-745-98	Sequence 98, Appl
26	48	49.5	807	12	US-10-063-512-98	Sequence 98, Appl
27	48	49.5	807	12	US-10-063-513-98	Sequence 98, Appl
28	48	49.5	807	12	US-10-063-515-98	Sequence 98, Appl
29	48	49.5	807	12	US-10-063-549-98	Sequence 98, Appl
30	48	49.5	807	12	US-10-063-569-98	Sequence 98, Appl
31	48	49.5	807	12	US-10-063-551-98	Sequence 98, Appl
32	48	49.5	807	12	US-10-006-485A-229	Sequence 229, App
33	48	49.5	807	12	US-10-013-907A-229	Sequence 229, App
34	48	49.5	807	12	US-10-015-499A-229	Sequence 229, App
35	48	49.5	807	12	US-10-063-555-98	Sequence 98, Appl
36	48	49.5	807	12	US-10-063-563-98	Sequence 98, Appl
37	48	49.5	807	12	US-10-063-553-98	Sequence 98, Appl
38	48	49.5	807	12	US-10-063-594-98	Sequence 98, Appl
39	48	49.5	807	12	US-10-063-554-98	Sequence 98, Appl
40	48	49.5	807	12	US-10-232-224-188	Sequence 188, App
41	48	49.5	807	12	US-10-013-910A-229	Sequence 229, App
42	48	49.5	807	12	US-10-226-254A-229	Sequence 229, App
43	48	49.5	807	12	US-10-015-395A-229	Sequence 229, App
44	48	49.5	807	13	US-10-006-867-98	Sequence 98, Appl
45	48	49.5	807	13	US-10-063-547-98	Sequence 98, Appl

ALIGNMENTS

RESULT 1

US-10-187-975-42

; Sequence 42, Application US/10187975

; Publication No. US20030224982A1

; GENERAL INFORMATION:

; APPLICANT: Li, Li

; APPLICANT: Shenoy, Suresh

; APPLICANT: Patturajan, Meera

; APPLICANT: Ellerman, Karen

; APPLICANT: Gorman, Linda

; APPLICANT: Zhong, Mei

; APPLICANT: Catterton, Elina

; APPLICANT: Spytek, Kimberly

; APPLICANT: Miller, Charles

; APPLICANT: Edinger, Shlomit

; APPLICANT: Hjalte, Tord

; APPLICANT: Gerlach, Valerie

; APPLICANT: Shinkets, Richard

; APPLICANT: Taupier, Raymond J. Jr.

; APPLICANT: Anderson, David

; APPLICANT: Guo, Xiaojia

; APPLICANT: Baumgartner, Jason

; APPLICANT: Padigaru, Muralidhara

; APPLICANT: Peyman, John

; APPLICANT: Smithson, Glennda

; APPLICANT: Casman, Stacie

; APPLICANT: Voss, Edward

; APPLICANT: Boldog, Ferenc

; APPLICANT: Pena, Carol

; APPLICANT: Chapoval, Andrei

; APPLICANT: Rastelli, Luca

; APPLICANT: Kekuda, Ramesh

; APPLICANT: Verite, Corine

; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING

; FILE OF INVENTION: SAME, AND METHODS OF USE

; FILE REFERENCE: 21402-397A

; CURRENT APPLICATION NUMBER: US/10/187,975

; CURRENT FILING DATE: 2002-07-02

;; PRIOR APPLICATION NUMBER: 60/303,046
;; PRIOR FILING DATE: 2001-07-05
;; PRIOR APPLICATION NUMBER: 60/303,828
;; PRIOR FILING DATE: 2001-07-09
;; PRIOR APPLICATION NUMBER: 60/304,502
;; PRIOR FILING DATE: 2001-07-11
;; PRIOR APPLICATION NUMBER: 60/305,011
;; PRIOR FILING DATE: 2001-07-12
;; PRIOR APPLICATION NUMBER: 60/305,262
;; PRIOR FILING DATE: 2001-07-13
;; PRIOR APPLICATION NUMBER: 60/305,673
;; PRIOR FILING DATE: 2001-07-16
;; PRIOR APPLICATION NUMBER: 60/306,085
;; PRIOR FILING DATE: 2001-07-17
;; PRIOR APPLICATION NUMBER: 60/307,536
;; PRIOR FILING DATE: 2002-07-24
;; PRIOR APPLICATION NUMBER: 60/308,228
;; PRIOR FILING DATE: 2001-07-27
;; PRIOR APPLICATION NUMBER: 60/308,877
;; PRIOR FILING DATE: 2001-07-30
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 288
;; SOFTWARE: CuraSeq1st version 0.1
;; SEQ ID NO 42
;; LENGTH: 999
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-187-975-42

Query Match 100.0%; Score 97; DB 12; Length 999;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EPNHLSKIAFKIVSQEPA 19
Db 186 EPNHLSKIAFKIVSQEPA 204
|||||

RESULT 2
US-10-157-031-26
; Sequence 26, Application US/10157031
; Publication No. US20030108890A1
; GENERAL INFORMATION:
; APPLICANT: Baranova, A. V.
; APPLICANT: Yankovsky, N. K.
; APPLICANT: Kozlov, A. P.
; APPLICANT: Lobashev, A. V.
; APPLICANT: Krukovskaya, L. L.
; TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequences
; FILE REFERENCE: 2760-103
; CURRENT APPLICATION NUMBER: US/10/157,031
; CURRENT FILING DATE: 2002-05-30
; NUMBER OF SEQ ID NOS: 415
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
; LENGTH: 999
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-157-031-26

Query Match 100.0%; Score 97; DB 14; Length 999;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EPNHLSKIAFKIVSQEPA 19
Db 186 EPNHLSKIAFKIVSQEPA 204
|||||

RESULT 3
US-10-295-027-322
; Sequence 322, Application US/10295027
; Publication No. US20030232350A1

;; GENERAL INFORMATION:
;; APPLICANT: Afar, Daniel
;; APPLICANT: Aziz, Natasha
;; APPLICANT: Ginsberg, Wendy M.
;; APPLICANT: Gish, Kurt C.
;; APPLICANT: Glynn, Richard
;; APPLICANT: Hevezi, Peter A.
;; APPLICANT: Mack, David H.
;; APPLICANT: Murray, Richard
;; APPLICANT: Watson, Susan R.
;; APPLICANT: Eos Biotechnology, Inc.
;; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
;; FILE REFERENCE: 018501-012500US
;; CURRENT APPLICATION NUMBER: US/10/295,027
;; CURRENT FILING DATE: 2002-11-13
;; PRIOR APPLICATION NUMBER: US 09/663,733
;; PRIOR FILING DATE: 2000-09-15
;; PRIOR APPLICATION NUMBER: US 60/350,666
;; PRIOR FILING DATE: 2001-11-13
;; PRIOR APPLICATION NUMBER: US 60/335,394
;; PRIOR FILING DATE: 2001-11-15
;; PRIOR APPLICATION NUMBER: US 60/332,464
;; PRIOR FILING DATE: 2001-11-21
;; PRIOR APPLICATION NUMBER: US 60/334,393
;; PRIOR FILING DATE: 2001-11-29
;; PRIOR APPLICATION NUMBER: US 60/340,376
;; PRIOR FILING DATE: 2001-12-14
;; PRIOR APPLICATION NUMBER: US 60/347,211
;; PRIOR FILING DATE: 2002-01-08
;; PRIOR APPLICATION NUMBER: US 60/347,349
;; PRIOR FILING DATE: 2002-01-10
;; PRIOR APPLICATION NUMBER: US 60/355,250
;; PRIOR FILING DATE: 2002-02-08
;; PRIOR APPLICATION NUMBER: US 60/356,714
;; PRIOR FILING DATE: 2002-02-13
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 1386
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 322
;; LENGTH: 999
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-295-027-322

Query Match 100.0%; Score 97; DB 15; Length 999;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 EPNHLSKIAFKIVSQEPA 19
Db 186 EPNHLSKIAFKIVSQEPA 204
|||||

RESULT 4
US-10-188-832-199
; Sequence 199, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder
; FILE REFERENCE: 018501-002330US
; CURRENT APPLICATION NUMBER: US/10/188,832
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/302,814
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/310,099
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/343,705

RESULT 9

```
US-10-006-869-32
; Sequence 32, Application US/10006869
; Publication No. US20030082166A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C7
; CURRENT APPLICATION NUMBER: US/10/006,869
; CURRENT FILING DATE: 2001-12-03
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-006-869-32

Query Match      73.2%; Score 71; DB 14; Length 113;
Best Local Similarity 73.7%; Pred. No. 0.00044;
Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EPNHLSKIAPKIVSQEPA 19
||| |||||:|||||
Db 29 EPNLNSKISYRIVSLEPA 47

RESULT 10
US-10-395-032-32
; Sequence 32, Application US/10395032
; Publication No. US20030229199A1
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C9
; CURRENT APPLICATION NUMBER: US/10/395,032
; CURRENT FILING DATE: 2003-03-21
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-395-032-32

Query Match      73.2%; Score 71; DB 15; Length 113;
Best Local Similarity 73.7%; Pred. No. 0.00044;
Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EPNHLSKIAPKIVSQEPA 19
||| |||||:|||||
Db 29 EPNLNSKISYRIVSLEPA 47

RESULT 11
US-09-909-567B-42
; Sequence 42, Application US/09909567B
; Publication No. US20030022257A1
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto A.
; APPLICANT: Nair, Manoj
; APPLICANT: Chen, Seiyu
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes
; FILE REFERENCE: DEX-0214
; CURRENT APPLICATION NUMBER: US/09/909,567B
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 60/219,834
; PRIOR FILING DATE: 2000-07-21
```

```
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 1118
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-909-567B-42

Query Match      73.2%; Score 71; DB 10; Length 1118;
Best Local Similarity 73.7%; Pred. No. 0.0058;
Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EPNHLSKIAPKIVSQEPA 19
||| |||||:|||||
Db 188 EPNLNSKISYRIVSLEPA 206

RESULT 12
US-09-746-491-43
; Sequence 43, Application US/09746491
; Patent No. US20020137202A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: NO. US20020137202A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-621
; CURRENT APPLICATION NUMBER: US/09/746,491
; CURRENT FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: USSN 60/171,329
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-746-491-43

Query Match      53.6%; Score 52; DB 9; Length 413;
Best Local Similarity 47.4%; Pred. No. 3.1;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 EPNHLSKIAPKIVSQEPA 19
:||| |||:|||||
Db 53 DPNALNGMLRYRIVSQAPS 71

RESULT 13
US-09-805-020-56
; Sequence 56, Application US/09805020
; Publication No. US20020086384A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, Zurit
; TITLE OF INVENTION: SPLICE VARIANTS OF ONCOGENES
; FILE REFERENCE: 2786-0168P
; CURRENT APPLICATION NUMBER: US/09/805,020
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56
; LENGTH: 848
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-805-020-56

Query Match      53.6%; Score 52; DB 12; Length 848;
Best Local Similarity 47.4%; Pred. No. 7;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 EPNHLSKIAPKIVSQEPA 19
:||| |||:|||||
Db 296 DPNALNGMLRYRIVSQAPS 314
```

```
RESULT 14
US-09-905-983-46
; Sequence 46, Application US/09905983
; Patent No. US20020045591A1
; GENERAL INFORMATION:
; APPLICANT: Geiger, Benjamin
; APPLICANT: Ben-Ze'ev, Avri
; APPLICANT: Sadot, Einat
; TITLE OF INVENTION: METHODS AND THERAPEUTIC COMPOSITIONS FOR TREATING CANCER
; FILE REFERENCE: 01/22326
; CURRENT APPLICATION NUMBER: US/09/905,983
; CURRENT FILING DATE: 2001-09-28
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 46
; LENGTH: 906
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-905-983-46

Query Match      53.6%; Score 52; DB 9; Length 906;
Best Local Similarity 47.4%; Pred.No. 7.5;
Matches      9; Conservative      5; Mismatches      5; Indels      0; Gaps      0;

QY      1 EPHNLNSKIAFKIVSQEPA 19
      :||| ||| : :||| |
Db      296 DPNALNGMLRYRIVSQAPS 314

RESULT 15
US-09-746-491-46
; Sequence 46, Application US/09746491
; Patent No. US20020137202A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: No. US20020137202A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-621
; CURRENT APPLICATION NUMBER: US/09/746,491
; CURRENT FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: USSN 60/171,329
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46
; LENGTH: 906
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-746-491-46

Query Match      53.6%; Score 52; DB 9; Length 906;
Best Local Similarity 47.4%; Pred.No. 7.5;
Matches      9; Conservative      5; Mismatches      5; Indels      0; Gaps      0;

QY      1 EPHNLNSKIAFKIVSQEPA 19
      :||| ||| : :||| |
Db      296 DPNALNGMLRYRIVSQAPS 314
```

Search completed: October 12, 2004, 06:06:11
Job time : 75 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2004, 05:58:17 ; Search time 22 Seconds
(without alignments)
44.586 Million cell updates/sec

Title: US-10-799-005A-1

Perfect score: 97

Sequence: 1 EPNHLNSKIAFKIVSQEPA 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/iaa/5A COMB.pep:*
- 2: /cgn2_6/ptodata/2/iaa/5B COMB.pep:*
- 3: /cgn2_6/ptodata/2/iaa/6A COMB.pep:*
- 4: /cgn2_6/ptodata/2/iaa/6B COMB.pep:*
- 5: /cgn2_6/ptodata/2/iaa/PTCUS COMB.pep:*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81	83.5	112	4	US-09-187-859-31
2	81	83.5	112	4	US-09-839-542B-31
3	81	83.5	112	4	US-09-535-852-31
4	71	73.2	15	1	US-08-787-547-24
5	71	73.2	15	2	US-08-400-756-3
6	71	73.2	113	4	US-09-187-859-32
7	71	73.2	113	4	US-09-839-542B-32
8	71	73.2	113	4	US-09-535-852-32
9	52	53.6	906	4	US-09-417-039-11
10	49	50.5	555	2	US-08-453-702A-98
11	49	50.5	556	1	US-07-998-003A-98
12	49	50.5	556	1	US-08-453-274B-98
13	49	50.5	556	1	US-08-453-695A-98
14	49	50.5	556	3	US-08-268-161A-98
15	49	50.5	556	3	US-09-099-639-98
16	45	46.4	619	4	US-09-578-441-3
17	45	46.4	620	4	US-09-578-441-4
18	45	46.4	884	2	US-08-474-067-8
19	45	46.4	884	2	US-08-474-068A-8
20	45	46.4	884	2	US-08-474-068A-7
21	43	44.3	913	2	US-08-472-481-7
22	43	44.3	913	2	US-08-474-067-6
23	43	44.3	913	2	US-08-474-068A-6
24	43	44.3	913	2	US-08-472-481-5
25	42	43.3	878	1	US-08-237-919-2
26	42	43.3	878	4	US-08-732-429-2
27	42	43.3	878	4	US-09-798-267-2
28	42	43.3	878	4	US-09-798-267-3

Sequence 2, Appli
Sequence 5068, Ap
Sequence 2, Appli
Sequence 2, Appli
Sequence 5, Appli
Sequence 2, Appli
Sequence 2, Appli
Sequence 17, Appli
Sequence 27, Appli
Sequence 27, Appli
Sequence 4196, Ap
Sequence 9, Appli
Sequence 9, Appli
Sequence 8, Appli
Sequence 13, Appli
Sequence 12, Appli
Sequence 30, Appli
Sequence 9125, Ap

28 42 43.3 878 5 PCT-US95-05518-2
29 41 42.3 208 4 US-09-134-001C-5068
30 41 42.3 626 1 US-07-938-782A-2
31 41 42.3 626 1 US-08-630-524-2
32 41 42.3 626 4 US-09-578-441-5
33 41 42.3 626 5 PCT-US93-08131-2
34 41 42.3 629 4 US-09-578-441-2
35 41 42.3 629 4 US-09-799-875-17
36 40 41.2 109 1 US-08-466-886-27
37 40 41.2 109 3 US-08-469-617-27
38 40 41.2 600 4 US-09-543-681A-4196
39 39 40.2 822 2 US-08-474-067-9
40 39 40.2 822 2 US-08-474-068A-9
41 39 40.2 822 2 US-08-472-481-8
42 38.5 39.7 113 2 US-08-676-782-13
43 38.5 39.7 124 2 US-08-676-782-12
44 38.5 39.7 370 4 US-09-331-568A-30
45 38.5 39.7 376 4 US-09-489-039A-9125

ALIGNMENTS

RESULT 1

US-09-187-859-31

; Sequence 31, Application US/09187859A

; Patent No. 6358920

; GENERAL INFORMATION:

; APPLICANT: Blaschuk, Orest W.

; APPLICANT: Gour, Barbara J.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL

; FILE REFERENCE: 100086.407C1

; CURRENT APPLICATION NUMBER: US/09/187,859A

; CURRENT FILING DATE: 1998-11-06

; NUMBER OF SEQ ID NOS: 4052

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 31

; LENGTH: 112

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-187-859-31

Query Match 83.5%; Score 81; DB 4; Length 112;
Best Local Similarity 78.9%; Pred. No. 6.6e-07;
Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EPNHLNSKIAFKIVSQEPA 19

Db 29 EPNHLNSKIAFKIIRQEPS 47

RESULT 2

US-09-839-542B-31

; Sequence 31, Application US/09839542B

; Patent No. 6569996

; GENERAL INFORMATION:

; APPLICANT: Blaschuk, Orest W.

; APPLICANT: Symonds, James Matthew

; APPLICANT: Gour, Barbara J.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL

; FILE REFERENCE: 100086.407D1

; CURRENT APPLICATION NUMBER: US/09/839,542B

; CURRENT FILING DATE: 2001-04-20

; NUMBER OF SEQ ID NOS: 4052

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 31

; LENGTH: 112

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-839-542B-31

Query Match 83.5%; Score 81; DB 4; Length 112;
Best Local Similarity 78.9%; Pred. No. 6.6e-07;
Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFKIVSQEPA 19
DB 29 EPNHLSKIAFKIIRQEPS 47

RESULT 3

US-09-535-852-31
; Sequence 31, Application US/09535852
; Patent No. 6638911
; GENERAL INFORMATION:
; APPLICANT: Blachuk, Orest W.
; APPLICANT: Symonds, James M.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND ETHODS FOR MODULATING
; TITLE OF INVENTION: DESMOSOMAL CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C6
; CURRENT APPLICATION NUMBER: US/09/535,852
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 2009
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-535-852-31

Query Match 83.5%; Score 81; DB 4; Length 112;
Best Local Similarity 78.9%; Pred. No. 6.6e-07;
Matches 15; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 EPNHLSKIAFKIVSQEPA 19
DB 29 EPNHLSKIAFKIIRQEPS 47

RESULT 4

US-08-787-547-24
; Sequence 24, Application US/08787547
; Patent No. 5783567
; GENERAL INFORMATION:
; APPLICANT: Hedley, Mary Lynne
; APPLICANT: Curley, Joanne M.
; APPLICANT: Langer, Robert S.
; TITLE OF INVENTION: MICROPARTICLES FOR DELIVERY
; TITLE OF INVENTION: OF NUCLEIC ACID
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/787,547
; FILING DATE: 22-JAN-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 08191/003001

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-787-547-24

Query Match 73.2%; Score 71; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVSQEPA 19
DB 1 LNSKIAFKIVSQEPA 15

RESULT 5

US-08-400-796-3
; Sequence 3, Application US/08400796
; Patent No. 5874531
; GENERAL INFORMATION:
; APPLICANT: STROMINGER, JACK L.
; APPLICANT: WUCHERPFENNIG, KAI
; TITLE OF INVENTION: IDENTIFICATION OF SELF AND NON-SELF
; TITLE OF INVENTION: ANTIGENS IMPLICATED IN AUTOIMMUNE DISEASE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/400,796
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: TWOMEY, MICHAEL J.
; REGISTRATION NUMBER: 38,349
; REFERENCE/DOCKET NUMBER: H0498/7015
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-720-3500
; TELEFAX: 617-720-2441
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: HOMO SAPIENS
US-08-400-796-3

Query Match 73.2%; Score 71; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 LNSKIAFKIVSQEPA 19
DB 1 LNSKIAFKIVSQEPA 15

```
RESULT 6
US-09-187-859-32
; Sequence 32, Application US/09187859A
; Patent No. 6358920
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C1
; CURRENT APPLICATION NUMBER: US/09/187,859A
; CURRENT FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-187-859-32

Query Match 73.2%; Score 71; DB 4; Length 113;
Best Local Similarity 73.7%; Pred. No. 3.9e-05;
Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EPNHLSKIAPKIVSQEPA 19
Db 29 EPTLNSKISYRIVSLEPA 47

RESULT 7
US-09-839-542B-32
; Sequence 32, Application US/09839542B
; Patent No. 6569996
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407D1
; CURRENT APPLICATION NUMBER: US/09/839,542B
; CURRENT FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-839-542B-32

Query Match 73.2%; Score 71; DB 4; Length 113;
Best Local Similarity 73.7%; Pred. No. 3.9e-05;
Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EPNHLSKIAPKIVSQEPA 19
Db 29 EPTLNSKISYRIVSLEPA 47

RESULT 8
US-09-535-852-32
; Sequence 32, Application US/09535852
; Patent No. 6638911
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James M.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING
; TITLE OF INVENTION: DESMOSONAL CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C6
; CURRENT APPLICATION NUMBER: US/09/535,852
```

```
; CURRENT FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 2009
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-535-852-32

Query Match 73.2%; Score 71; DB 4; Length 113;
Best Local Similarity 73.7%; Pred. No. 3.9e-05;
Matches 14; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EPNHLSKIAPKIVSQEPA 19
Db 29 EPTLNSKISYRIVSLEPA 47

RESULT 9
US-09-417-039-11
; Sequence 11, Application US/09417039A
; Patent No. 6485972
; GENERAL INFORMATION:
; APPLICANT: McMahon, Andrew P.
; APPLICANT: Parr, Brian A.
; APPLICANT: Vaino, Seppo
; TITLE OF INVENTION: WNT SIGNALING IN REPRODUCTIVE ORGANS
; FILE REFERENCE: 00246/232001
; CURRENT APPLICATION NUMBER: US/09/417,039A
; CURRENT FILING DATE: 1999-10-12
; EARLIER APPLICATION NUMBER: US 60/109,355
; EARLIER FILING DATE: 1998-10-15
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 906
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-417-039-11

Query Match 53.6%; Score 52; DB 4; Length 906;
Best Local Similarity 47.4%; Pred. No. 1;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 1 EPNHLSKIAPKIVSQEPA 19
Db 296 DPNALNGMLRYRIVSQAPS 314

RESULT 10
US-08-453-702A-98
; Sequence 98, Application US/08453702A
; Patent No. 5891706
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: Protocadherin Materials and Methods
; NUMBER OF SEQUENCES: 115
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
; ADDRESSEE: Borun
; STREET: 233 South Wacker, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,702A
; FILING DATE:
```

```

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5891706and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 32657
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 555 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-453-702A-98

```

```

Query Match 50.5%; Score 49; DB 2; Length 555;
Best Local Similarity 42.1%; Pred. No. 1.9;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 EPNHNSKIAFKIVSQEPA 19
:||||| : : : : :
Db 137 DPNALNGMLRYRILSQAPS 155

```

```

RESULT 11
US-07-998-003A-98
; Sequence 98, Application US/07998003A
; Patent No. 5643781
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: Protocadherin Materials and Methods
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
; ADDRESSEE: Bicknell
; STREET: 20 South Clark Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/998,003A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5643781and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 30903
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/346-5750
; TELEFAX: 312/984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 556 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-998-003A-98

```

```

Query Match 50.5%; Score 49; DB 1; Length 556;
Best Local Similarity 42.1%; Pred. No. 1.9;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 EPNHNSKIAFKIVSQEPA 19
:||||| : : : : :

```

```

Db 137 DPNALNGMLRYRILSQAPS 155
RESULT 12
US-08-453-274B-98
; Sequence 98, Application US/08453274B
; Patent No. 5663300
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: Protocadherin Materials and Methods
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,274B
; FILING DATE: 30-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5663300and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 32660
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 556 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-453-274B-98

```

```

Query Match 50.5%; Score 49; DB 1; Length 556;
Best Local Similarity 42.1%; Pred. No. 1.9;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 EPNHNSKIAFKIVSQEPA 19
:||||| : : : : :
Db 137 DPNALNGMLRYRILSQAPS 155

```

```

RESULT 13
US-08-453-695A-98
; Sequence 98, Application US/08453695A
; Patent No. 5708143
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: Protocadherin Materials and Methods
; NUMBER OF SEQUENCES: 115
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
; ADDRESSEE: Borun
; STREET: 233 South Wacker, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/453,695A
FILING DATE: 530
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: No. 5708143and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 32658
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 98:
SEQUENCE CHARACTERISTICS:
LENGTH: 556 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-453-695A-98

Query Match 50.5%; Score 49; DB 1; Length 556;
Best Local Similarity 42.1%; Pred. No. 1.9;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 1 EPNHLSKIAPKIVSQEPA 19
Db 137 DPNALNGMLRYRLSQAPS 155

RESULT 14

US-08-268-161A-98
Sequence 98, Application US/08268161A
Patent No. 5798224
GENERAL INFORMATION:
APPLICANT: Suzuki, Shintaro
TITLE OF INVENTION: Protocadherin Materials and Methods
NUMBER OF SEQUENCES: 115
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
ADDRESSEE: Borun
STREET: 233 South Wacker, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/268,161A
FILING DATE: June 27, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Young J. Suh
REGISTRATION NUMBER: P-41,337
REFERENCE/DOCKET NUMBER: 27866/32149
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 98:
SEQUENCE CHARACTERISTICS:
LENGTH: 556 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-268-161A-98

Query Match 50.5%; Score 49; DB 1; Length 556;
Best Local Similarity 42.1%; Pred. No. 1.9;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 1 EPNHLSKIAPKIVSQEPA 19
Db 137 DPNALNGMLRYRLSQAPS 155

RESULT 15

US-09-099-639-98
Sequence 98, Application US/09099639
Patent No. 6262237
GENERAL INFORMATION:
APPLICANT: Suzuki, Shintaro
TITLE OF INVENTION: Protocadherin Materials and Methods
NUMBER OF SEQUENCES: 115
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray, &
STREET: 233 South Wacker, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/099,639
FILING DATE: 18 JUN 1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/263,161
FILING DATE: 27 JUN 1994
ATTORNEY/AGENT INFORMATION:
NAME: Greta E. No. 626237and
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27866/34703
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 98:
SEQUENCE CHARACTERISTICS:
LENGTH: 556 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-099-639-98

Query Match 50.5%; Score 49; DB 3; Length 556;
Best Local Similarity 42.1%; Pred. No. 1.9;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 1 EPNHLSKIAPKIVSQEPA 19
Db 137 DPNALNGMLRYRLSQAPS 155

Search completed: October 12, 2004, 06:04:52
Job time : 24 secs

This Page Blank (uspto)